

18 June 1981

Making Israel atone for Tamuz

The chief casualty of the Israeli raid on the Iraqi reactor complex at Tamuz is the international reputation of the Non-Proliferation Treaty. Iraq, a ratified signatory of the treaty since 1969, in 1977 agreed with the International Atomic Energy Agency that all its nuclear operations should be open to formal inspection. What Israel has now done is to demonstrate to Iraq, and to other Arab and Muslim states, that formal obligations such as these provide no protection from Israeli interdiction. Even the government of Egypt, for the time being in Mr Menachem Begin's good books, must be wondering whether it will be immune from a similar attack when the power reactor to be built under an agreement with the United States, again under the aegis of the Non-Proliferation Treaty, is nearing completion. If the present Israeli mood survives the coming general election, the agreements carefully negotiated between Israel and Egypt could be put on the back burner for the sake of a pre-emptive strike against Egyptian competence in civil nuclear technology. Before that happens, and for the sake of repairing the political damage already done in the Middle East, the interested parties, but especially the governments of France and the United States, should look for a remedy that will strengthen the Non-Proliferation Treaty. Is it too much to ask that Israel itself should now be made to sign the treaty?

Israeli policy on nuclear technology in the past twenty years is not at first sight encouraging, yet the events of the past two weeks are an opportunity as well as a disaster. Ever since the construction of the Israeli reactor at Dimona in the Negev in the early 1960s, Israel has enjoyed the benefits of being a putative nuclear power. Israel's neighbours have been encouraged to suppose that Israel has a modest stock of nuclear weapons that could and would be used in anger in an emergency. This reputation has served Israel well. Some neighbours of Israel, Egypt and Syria in particular, have no doubt had to take into account the possibility of nuclear retaliation when there has been trouble in the Sinai or on the Golan Heights. On the other hand, Israel has escaped the contumely that would follow a tangible demonstration that it has nuclear weapons at its disposal. As things are, nobody knows for sure whether Israel has nuclear weapons or not — the only compensation is that if there are Israeli bombs, Israeli generals must be less confident than their opposite numbers elsewhere that they will function as intended. What are the chances that in the aftermath of the Tamuz raid, it will be possible to force Israel to come clean?

The Israeli justification of the raid on Iraq is paper-thin. That the raid was an open breach of international law is not disputed. That would have been the case even if, as the government of Israel claims, there were evidence that Iraq had been planning to use the 70 megawatt reactor being built under its agreement with France as a source of military plutonium. But in reality, whatever dreams the reactor managers of Tamuz may have been cherishing, the simple truth is that Iraq could not have made weapons with plutonium from its new reactor without the International Atomic Energy Agency knowing well in advance. The agreement between Iraq and France requires that enriched uranium for the reactor should be returned to France for reprocessing, while an attempt to generate plutonium from natural uranium (supposedly bought from Portugal) would surely have been spotted by the least competent of the inspectors from Vienna. Only in two respects could the developments in Iraq have been a military threat to Israel. Iraq might at some point in the future have abrogated its adherence to the Non-Proliferation Treaty, and set about using the reactor as a source of military plutonium. So far, no state has

taken such a step, and while the major powers remain united behind the treaty none is likely to dare. The result of abrogation would be intolerable international pressure, the risk of not surviving the period between abrogation and becoming a nuclear power — and most probably a more legitimate version of last week's Israeli raid. Alternatively, Iraqi expertise might have been used as a way of helping like-thinking states, not signatories of the Non-Proliferation Treaty, to make nuclear weapons. Pakistan is most often mentioned. Strictly, assistance of this kind is against the rules. In practice, violations are harder to detect or at least to prove. But again, there is no reason why the government of Israel should have taken it on itself to pre-empt such developments.

But why should Iraq be interested in nuclear power? Should not oil-rich states content themselves with drilling holes in the ground? That is a second line of defence of the Israeli raid advanced by Mr Begin and his apologists elsewhere. The simple answer is "Why not?" Iraq is not, in any case, the first oil-rich state to take an interest in civil nuclear power. Iran in the Shah's day was an apparently enthusiastic buyer of nuclear reactors from France and the United States. The logic was simple. Why use chemically sophisticated hydrocarbons as a source of heat when they could be used as starting points for chemical manufacturing instead? More crudely, why not use what seemed once to be cheap nuclear energy to generate electricity, allowing large amounts of oil to be sold at \$40 a barrel or whatever it will make on a seller's market? There is no reason why Iraq should not (at least in 1977) have nursed such an ambition. It is in any case as plausible that Iraq may have been calculating — that the calculation may have been wrong is irrelevant — that a dash of civil nuclear technology might be good for a technologically backward state. That prescription has worked to some incalculable effect in India. And if it is true that Iraq and France have been less open about their agreement on the Tamuz reactors (there is — or was — also a 600 kilowatt reactor) than they should have been, who, in Israel at least, can complain if the parties to the agreement kept it to themselves for fear of provoking trouble from Israel?

For the time being, responsibility for responding to the Israeli raid rests with the governments of Iraq, France and the United States, in that order of immediacy. What should they do? Iraq, laughably enough, probably has a civil suit against Israel; in spite of the cumbersomeness of international legal processes, that should be pursued. There is no reason why the reactor at Tamuz should not in due course be replaced, and at somebody else's expense. For the new government of France, the raid on Tamuz could hardly have been more ineptly timed. The new president had plainly been looking forward to reconstructing French relationships in the Middle East, perhaps resurrecting some of the friendliness between France and Israel that ended disastrously at Suez in 1957; he will now have been given pause. There is, however, no reason to suppose that the new French government will be as enamoured as was M. Jacques Chirac, the Prime Minister of France in 1977, of assisting with the creation of independent nuclear enterprises elsewhere. The chances are that President Mitterrand will confine himself, for the time being, to the moderate protests of the past few days. Further ahead, however, there is at least a chance that the new France may be tempted to sign the Non-Proliferation Treaty. M. Mitterrand has cautiously avoided all but passing references to the subject. Somebody should ask him what is in his mind.

The principal burden will, however, rest with the United States, hamstrung as always by what is called the Zionist vote. President



Reagan has found a neat way of shifting the immediate burden of awkward choice from the White House to the Senate Foreign Relations Committee, which will have to confirm or deny his suspension of aircraft sales to Israel. Sooner or later, however, the buck will come to rest again in the Oval Office. At some point, the Administration will have to devise a new anti-proliferation policy — the Carter version is too inflexible, and ideologically unsuited to this Administration. With the Tamuz raid fresh in people's minds, Mr Reagan will be repeatedly reminded (by Arab states if not by his voters) that the aircraft which carried out the raid on Tamuz were supplied by the United States and would have suited well for the delivery of nuclear weapons. Does it make sense, people will be asking, that the United States should resist the proliferation of nuclear weapons but be prepared to supply

delivery systems to putative nuclear powers such as Israel? And how can the United States, which has invested a vast amount of political capital in the past few years in the good cause of persuading other states to sign the Non-Proliferation Treaty, now give Israel (which has not signed) a licence to undercut that policy? It is not as if Israel's supporters, in the United States and elsewhere, seriously consider that Israel's safety depends on the undefined threat of retaliation with nuclear weapons. The supporters of Israel's continued independence would indeed be more robust — and more numerous — if there were not the continually nagging fear that help would be used eccentrically, even illegally. This is why the time has come for Israel to sign the Non-Proliferation Treaty — and for the United States to ensure that it is done.

Cheque-book academics in trouble

Stanford University seems ready to take a cautious step towards an accommodation between the principles of the academic life and the wish of some academics to turn their expertise to profit. The proposals (see page 526), on which the Stanford faculty will be asked to decide in a few months' time, have been made necessary by the apparently ceaseless spawning of commercial companies aiming to exploit genetic manipulation for commercial purposes and which include academic researchers as stockholders or even as entrepreneurs. The objective is to avoid as far as possible the conflicts of interest that arise when academics individually, or a university corporately, become dependent on commercial enterprises. The method is simple enough, almost deceptively so. The university will not (if the proposals are accepted) take an equity interest in commercial enterprises in which members of the faculty have a similar interest, while individual teachers will be required to disclose their outside interest whenever asked to do so by some more senior person, a dean or even the president of the university.

The university's proposed act of self-denial is virtually a replica of the decision reached at Harvard last October, when the faculty there persuaded the university administration not to go ahead with a proposal to accept (without payment) a block of shares in a genetic engineering company being organized by Professor Mark Ptashne. The arguments then, as now at Stanford, were straightforward. With a substantial minority stake in a successful business, a university may find it hard to avoid giving those who are nurturing the golden goose unfair advantages (compared with colleagues) in matters such as appointments or even the allocation of space. This, however, is the least of the worries likely to confront universities with equity stakes in companies. What, for example, would happen if such a company were in financial trouble? Would it be possible for a university to resist pleas for just a little help, a little cash perhaps, or forbearance with bills inconveniently receivable? Harvard and now Stanford are right to decide that the risks are too great too be stomachied. Neither university has so far taken a view on the propriety of the role of academics as spare-time entrepreneurs.

The obvious reason is that most university faculties would be hopelessly divided on any such proposition. So much is clear from the way in which Stanford has sidled up to the proposition — not yet agreed — that members of the faculty should disclose their outside interests to the academic administration whenever it seems possible that a conflict of interest has arisen. Why wait until then? Why not simply agree that academics have a duty to disclose all their business interests to their immediate colleagues, not merely to august administrators with an ear for gossip? The obvious drawback in the arrangement now proposed at Stanford is that academics with outside interests (not necessarily commercial) are not the best placed to decide that a conflict has arisen with their academic responsibilities. Colleagues are likely to be more sensitive. They should be given the right to express reasonable doubts, no doubt within a framework of understanding that the process of innovation is often furthered by

means of links, consultancy agreements for example, between academics and industry.

Stanford, then, should opt for full disclosure, at the level of the department, of academics' outside interests. The outcome would probably be surprisingly dull. Most people, it would emerge, have modest consultancy arrangements with the odd company or two, consider themselves ill-rewarded for the advice they give and make up for that by skimping on what they do. Only occasionally would it emerge that an academic's commitment to some commercial company was offensive, perhaps because of being exclusive to one external organization, perhaps (and worse) because it endangered the proper balance of a research programme or the education of graduate students.

What, against this background, is to be made of the growing company of academic entrepreneurs, the people who help to start commercial companies, often by telling the tale of the promise hidden in their research to the moneybags who must in the end provide what is called the "venture capital"? The fashion, in the past few months, for taking on this role has been most conspicuous in the application of molecular biology but is unlikely to be thus restricted for very long. Now that molecular biologists have shown how it is possible to make paper fortunes when a go-go enterprise is launched on an unsuspecting stock market, it is unlikely that computer scientists, organic chemists and people in general will hang back. The simple truth, however, is that there is a difference between a consultant — a *per diem* person — and one who helps to start a company. The former can walk away from disaster with an easy conscience; the latter must redouble his efforts when disaster threatens. The conclusion must surely be that an academic who opts for the role of entrepreneur must at the same time opt out of tenure. In the last resort, the argument must go, the faculty cannot reasonably command such a person's loyalty.

In the United States, the seemingly regulation of these questions would be simpler if it had not in the past few years been complicated by the amendments of the Patents and Trademarks Act. In the old days, the outcome of a federally-sponsored research programme was in the public domain and thus up for grabs from anybody with the wit to recognize its importance. Now, universities at which research projects are undertaken have a right, even a responsibility, to patent what they can and to exploit it commercially. The result is often a nonsense, tacitly acknowledged as such in the burgeoning field of monoclonal antibodies (see page 525). The snag is that universities by their constitution are not much good at playing patent agents, but (given the continuing threat of shrinking funds) awkwardly eager to do so. The message seems to have spread quickly to teaching staffs. All concerned (the federal agencies in the United States included) should brood on what they have done before things go too much further. Elsewhere, in Britain for example, academics will soon be envying what seems to be the luxury of making money decisions. But worrying about earning money can be as sapping of scholarship as knowing that there is none.

Hoechst makes deal with Mass. General

House panel examines overseas ties

Washington

University involvement with the commercial exploitation of genetic engineering is well on the way to becoming an issue between the United States government and the universities. Last week, at a meeting of Congressman Albert Gore's oversight subcommittee of the House Committee on Science and Technology, representatives of the Massachusetts General Hospital were hauled over the coals because of the hospital's agreement with Hoechst AG under the terms of which the German chemical company will provide \$50 million over the next ten years to run a molecular biology laboratory at the hospital, and will have first refusal of any patent licences that may result.

At the hearings, Dr Ronald Lamont Havers, Director of research at the hospital, demurred when asked to provide the subcommittee with a copy of the agreement it has signed with Hoechst. The subcommittee's chairman left the hospital and everybody else within earshot in no doubt of his belief that universities sustained for the past twenty years by federal grants should be careful not to give overseas enterprises favoured access to their expertise. There is now some talk that the subcommittee will subpoena the agreement between the Massachusetts General Hospital and Hoechst if the document does not turn up in the mail before too long.

The grant from Hoechst will be used to found a department of molecular biology at the hospital. Dr Howard M. Goodman of the University of California, San Francisco, will be the director of the enterprise, the staff of which is expected to grow to about 50 by 1983 and thereafter to 100. The hospital insists that the programme of research will be determined by the hospital alone, while the work of the department will be reviewed by a committee of six scientists independent of Hoechst. The operation of the agreement between the company and the hospital will be kept under review by a joint committee of three senior managers of the company and three trustees from the hospital.

The declared objective of the new department is to apply the techniques of molecular biology to the treatment of disease. The hospital says that all appointments in the new department will be made according to established academic procedures, that individual scientists will be free to publish how and when they choose

provided that the hospital authorities are informed in advance, and that there will be no restrictions on collaboration with scientists elsewhere.

According to the hospital, patentable discoveries will be patented by the hospital, and the benefits shared between the hospital and the inventor on the basis of existing rules. Hoechst will have a right to a licence to any patent springing from research which it has sponsored, while the royalty rate negotiated "will reflect the financial contribution of the company".

Two potentially contentious aspects of the agreement are that the company will be free to decide whether or not particular

research projects should be funded out of the \$50 million set aside, and that members of the new department working on Hoechst-sponsored projects will not be free to consult with other companies. The agreement is automatically renewable if not terminated after ten years.

It has also been agreed that there will be a public seminar once a year to which some Hoechst scientists will be invited, and that the company will have a right to send up to four of its people to the hospital for training at any time.

Congressman Gore is unlikely to let these issues fade away. After last week's hearings, he said that "the questions that

NIH plan new overhead calculation

Academic research administrators will soon have to learn new algorithms for calculating how much overhead to charge against successful applications for research grants and contracts with the National Institutes of Health (NIH). So much is clear from the meeting last week (8-9 June) of the Institute's Advisory Committee, at which the task force led by Professor Samuel O. Thier of Yale University that has been brooding on the question of "institutional support" (bureaucratic jargon for overhead) promised to produce at the next meeting in October several proposals for reorganizing the present system.

Objections to the present system are several. Research grant proposals approved by study sections of NIH (the "peer review" committees) are made more costly by an amount negotiated between NIH and the recipient institutions intended to cover the cost of supporting the research concerned. On the average, NIH overheads cost 30 per cent of total expenditure on research grants and contracts. There are wide geographical variations, with the universities in the north-west of the United States successfully claiming larger percentages for institutional support than universities elsewhere.

The stimulus for the present review has come from federal agencies and universities, both equally appalled at the difficulties of carrying out the detailed accounting for research grant expenditure now required of them. But NIH also have an interest in heading off trouble about institutional support from the Reagan Administration, which has already shown its hand by removing from the NIH budget institutional support accompanying awards of postdoctoral fellowships.

The Thier task force plans to spend the summer assessing the merits of alternative mechanisms for providing institutional support. The favoured alter-

native to the present system appears to be the "fixed obligation" formula, under which research grants and contracts would be awarded on what is essentially a fixed-price basis, and on which federal investigation would be limited to a simple verification that direct and indirect costs have been properly incurred.

At the next meeting of the Advisory Committee in October, the task force will suggest that experiments should be carried out with several of the alternative methods of financial research projects. It is, however, unlikely that all the loose ends can be tied up by October — it is not, for example, at this stage clear whether the overhead element in fixed obligation grants should be assessed by a peer review committee and, if so, whether should be the appropriate study section.

Last week's meeting of the NIH Advisory Committee also wrestled a little inconclusively with the issue of patent rights in biological innovations.

One curious development to come to light was that many scientists are cheerfully ignoring the provisions of the Patents and Trademark (Amendments) Act of 1979. This legislation gives universities at which research grants are held the right to patent and exploit new developments, reserving to the federal government a non-exclusive right to a licence. Under the amended law, those holding grants from NIH are required to report all patentable developments.

These provisions have already been found unworkable in the development of monoclonal antibodies, each one of which may be potentially patentable. Given that a productive laboratory may expect to develop several hundreds of monoclonal antibodies a year, and that the cost of patent protection is a minimum of \$2,000, it seems to have been tacitly assumed that only those monoclonal antibodies likely to be commercially important deserve the investment of \$2,000-plus.

will confront our civilization as a result of genetic engineering?" will be the most difficult ever, and that their solution will only be complicated if those responsible for original discoveries make commercial links with industry.

The subcommittee last week found a more congenial witness in Dr Donald Kennedy, president of Stanford University, who earlier this year was advocating a "second Asilomar conference" on the problems of university-industry relationships. Kennedy, who more recently has been back-peddalling on this proposal, told the subcommittee last week of the ambivalence of research universities towards proposals for industrial support in the present climate of stagnant federal budgets. Well-ordered agreements, he said, could be beneficial, but he also pointed to the dangers of conflicts of interest, secrecy and the distortion of research training.

Universities are perhaps more than usually aware of these problems because of the advocacy of the virtues of industrial support for the research universities by the National Science Board.

Exploiting research

Stanford rules

Washington

A proposal for regulating the commercial activities of members of the Stanford University faculty is to be put to the Academic Senate at its meeting in September. One of the proposals is that members of the faculty should disclose details of consultancy agreements with commercial organizations if there are reasons to expect that conflicts of interest have arisen.

The new scheme has been worked out by a committee under Dr Ingram Olkin, professor of statistics and education at Stanford, which has recommended that the university should not invest in commercial ventures set up to exploit the results of research carried out at Stanford "if any current Stanford faculty member participates" either as a shareholder or as a manager. Similarly, the university would not accept stock in such a company as part of a licensing agreement.

The proposals have been stimulated by the recent interest in the commercial exploitation of genetic manipulation, to which several members of the Stanford faculty are in a position to contribute. But it is acknowledged that similar problems may arise in other fields, the development of computer software for example.

The case for a distant relationship between the university and commercial enterprises in which faculty members are involved is based on the fear of the conflicts of interest that could arise in making faculty appointments and promotions, allocating space and settling faculty salaries, the danger that graduate

Yugoslavs to resign

Dr Pavle Savic, President of the Serbian Academy of Arts and Sciences since 1971, has resigned in protest against the conduct of recent elections to the Academy. His resignation, which takes effect from the academy meeting of 28 May, came as no surprise — it had been openly discussed in the Yugoslav press for the preceding three weeks.

According to Dr Savic, he resigned because of the election of someone who had served several years' rigorous imprisonment as a "cominformist". He had been informed of this one month before the elections. On checking the allegation and finding it true, Dr Savic notified the appropriate Academy officials. But the election went ahead and the "cominformist" was elected. The next day, Dr Savic announced his intention to resign. Two vice-presidents and the general secretary of the academy also resigned; they, however, will serve out their full term of office (until the end of the year) to maintain continuity of the academy's activities.

Dr Savic's walk-out, seven months before his term of office ended, is widely seen as a protest, not merely against the election, but against the prevailing atmosphere in the academy. A physicist, he had hoped, to build up the research side of the academy, concentrating on modern trends including laser physics, solid-state electronics and molecular biology. He was particularly keen to bring in younger scientists into the academy's institute. Most members, however, showed little enthusiasm for his plans.

Vera Rich

education might be distorted and the possibility that the university's external reputation might be affected adversely.

The committee also advocates a change in the method by which the rewards from agreements for the licensing of patents are shared out. At present, if Stanford research leads to a patent, the income is divided equally between the researcher, his or her department and the university as a whole. It is now suggested that if the income from some invention should exceed one half of a department's annual budget, the two-thirds of the royalty income not due to the inventor should be paid into a research fund from which all departments at Stanford could hope to benefit.

The committee has obviously been cautious, perhaps even overcautious, in its approach to consultancy agreements. At present, academics are required to disclose the amount of time they spend in this way, which must not exceed 13 days a quarter. There is no mechanism for monitoring this gentleman's understanding. In future, if the faculty agrees, faculty members may have to disclose who they work for, and for how long.

Tropical disease finance

World Bank acts

After a plea by Mr Robert McNamara, who resigns as president of the World Bank on 30 June, the bank governors are expected to agree this week to contribute \$2.48 million (£1.25 million) to the Special Programme for Research and Training in Tropical Diseases (TDR) of the World Health Organization (WHO).

Previously, the bank has been willing to act merely as a financial adviser and banker to TDR, as one of the programme's three co-sponsors. (The others are WHO and the United Nations Development Programme.) The bank's main concern is with development loans, and it has been exploring for more than a year how to refine its policy towards the direct support of research. That review is not complete, but McNamara thought it well enough advanced that a commitment should now be made to the tropical disease programme. He carried his executive board with him, and their recommendation has now to be approved by the governors — the nations which fund the bank. There is not expected to be any opposition.

The support for TDR comes at an opportune time. First firmly established in 1977, and dealing with six diseases — malaria, schistosomiasis (bilharzia), trypanosomiasis (sleeping sickness), leishmaniasis, leprosy and filariasis — TDR's funds first grew rapidly but have now become roughly static at about \$24 million (£12 million) a year. Moreover Britain, which was contributing around £600,000 a year (but actually received back more in research grants, according to WHO), announced last December that it could not pay its 1981 contribution — shocking the 25 other contributing governments, many of which see little or no return in terms of research carried out in their countries.

According to the Canadian chairman of the TDR Scientific and Technical Advisory Committee, Dr A. B. Morrison, the effectiveness of the programme could be "crippled in the short term" if inflationary pressures are not compensated for; and in the long term, additional resources must be mobilized to permit large-scale trials of new tools for prevention and treatment now being developed.

Even so, Britain's contribution to TDR for the present year is definitely cancelled, says the Overseas Development Administration (ODA). And although next year's is "under review", the ODA's £1,000 million of aid money this year will fall by 5 per cent a year for the next three years, making a recovery of the TDR contribution less and less likely. This should not be taken as an implicit criticism of TDR, say ODA officials. On the contrary, the programme is the best managed of all the WHO special programmes, they say, thus putting a cloud over the two remaining programmes still

supported by ODA.

British policy is to cut multilateral, special programme aid hardest, for the government's aim is to maintain direct core subscriptions to international agencies (such as the World Bank and the United Nations Development Programme), while stressing bilateral aid which is inevitably under closer British political and economic control.

Robert Walgate

Armadillos fight leprosy

The safety and ethics panel of the World Health Organization (WHO) is likely to agree within the next few months on immunization trials with a new leprosy vaccine made from the liver and spleen of infected armadillos. The trials would be conducted from January to September 1982 in the United States, United Kingdom and Scandinavia to avoid criticisms of uncontrolled drug testing in third world countries.

The objective will be to induce 85–90 per cent immunity to the causative agent of leprosy (*Mycobacterium leprae*) as judged by immune reactions to lepromin, a preparation of killed *M. leprae* obtained from infected human ear-lobes.

The vaccine has been a long-standing objective of the WHO Special Programme for Research and Training in Tropical Diseases (TDR). The snag is that it depends on culturing the bug in armadillos, whose low body temperature allows growth of the organisms, and some have questioned whether it will ever be possible to breed enough armadillos to vaccinate the 2,000 million or so of the human population potentially at risk.

Dr Barry Blum of the Albert Einstein College of Medicine in New York, who heads the TDR leprosy immunization project, says, however, that armadillos are such prolific generators of *M. leprae* that the numbers should be sufficient. A mixed vaccine made from human tissue yields immunity with a single dose of 6×10^8 killed bacteria. An armadillo three years after infection produces about 2.5×10^{12} bacteria in the liver and spleen. If the vaccine produced from them is as effective as the human one, each armadillo should yield 4,000 doses. WHO now has 250 animals used for research and the forthcoming trial, but there is a population of 10 million armadillos in the southern United States.

The adequacy of the supply depends directly on the dose required to produce immunity. One shot must be 85–90 per cent effective to avoid the need for repeated injections. The availability of the vaccine would also depend on the efficiency of extraction from armadillo tissue, a problem which has been given to the Wellcome Laboratories in the United Kingdom to resolve. Then leprosy is such a slow-moving disease that proof in the field against infective organisms will take ten years.

Robert Walgate

US solar energy

No silver lining

Washington

The Solar Energy Research Institute at Golden, Colorado, has been flung further into limbo by the arrival of the new Administration, which is now trying to respond to a demand from the General Accounting Office that the objectives of the institute should be spelled out. Most probably the institute, once the spearhead of President Carter's ambition that a fifth of United States energy consumption at the end of the century should be solar energy, will be asked to concentrate on long-term research, abandoning the commercial demonstration of solar energy devices. The present budget of \$130 million is likely to be cut by a half, and between 100 and 200 of the present staff of 770 will probably lose their jobs.

The brief history of the institute, set up in 1977 and managed for the Department of Energy by the Mid-West Research Institute of Kansas, is a tale of muddle. Originally, there was to have been a staff of 1,200 engaged on long-term research, development and the commercial demonstration of devices able to win energy from the Sun. In part, the institute has been the victim of the success of some applications of solar energy. Solar water heaters, large and small, and small hydroelectric plants are multiplying throughout the United States — the Department of Energy is embarrassed by more than 1,000 applications from private companies and individuals to build hydroelectric plants. With the Administration committed to the view that commercial developments should be the responsibility of commercial organizations, not the government, the institute's demonstration programme has become vulnerable. So much seemed clear during a visit to Golden by officials of the Department of Energy last month.

Even when the axe falls, in the next few weeks, the rump of the institute is unlikely to be reassured about its future. Although the institute is now nearly four years old, the special building (on a 300-acre site offered by the Government of Colorado) seems as far off as ever. In a sharp letter to the Secretary of Energy at the end of April, the General Accounting Office asked that further design should be postponed until the department had decided what the long-term role of the institute should be, and that even the use of the 300-acre site for experimental rigs should wait on this definition. One snag is that the offer of the site expires in April 1982.

The letter is scathing about the planning of the building. Although authority for a laboratory was included in the 1974 Solar Energy Act, the estimated cost of \$132 million (in 1978) was an unwelcome surprise. Successive Secretaries of Energy have sought to limit the cost of the buildings to \$98 million (in 1979) and \$75 million

(in 1980). The General Accounting Office points out that the original scheme for making the building 80 per cent self-sufficient in its own energy requirements was based on the use of active and passive solar power devices which had not been proved and which were not all cost-effective. Meanwhile the institute's staff camps out in rented office accommodation costing \$5 million a year, and has spent substantial sums on the conversion of offices to laboratories.

Since 1977, the institute's research programme has been trimmed by the removal of general responsibility for the exploitation of biomass, but it remains responsible for the use of methanol as a liquid fuel — a motor vehicle driven by hydrogen derived from the catalytic decomposition of methanol will be demonstrated later in the year. The institute is also confident that it will be able to produce silicon-based solar cells yielding power at a cost of 70 cents per watt by the end of this decade, roughly a tenth of the present cost of solar cells.

European Space Agency

Delays in space

For the second time this year, the European Space Agency (ESA) may be hamstrung by a change of plans within the United States National Aeronautics and Space Administration (NASA). This time, the possible casualty is Spacelab, ESA's manned space laboratory due to be launched on the space shuttle. Spacelab's first flight could be delayed, or its success reduced, if the Department of Defense persuades NASA to modify and thus delay the launch of the second of the two satellites intended to form the basis of a new data collection system.

The Tracking and Data Relay Satellite System (TDRSS) will eventually replace much of the worldwide network of tracking stations for relaying data from the space shuttle and satellites in low Earth orbit. Two tracking satellites, to be launched into geosynchronous orbits from early shuttle flights, were to transmit data from almost any position in a low Earth orbit to a central receiving station at White Sands, New Mexico. The original plan was to launch the first tracking satellite on the sixth shuttle flight, the second in June 1983 and Spacelab itself in September 1983. But now the Department of Defense has asked that extra coding should be built into the second satellite which could delay the launching until after that of Spacelab. One possibility is that the first Spacelab flight would be brought forward to June 1983 and the second tracking satellite launched only in September 1983.

If NASA accedes to this request, ESA would have the choice of either flying the first Spacelab payload with only one tracking satellite in service or of delaying the flight until the spring of 1984. One

All aboard for Halley

The planned Soviet-French joint mission to Halley's comet in 1986 has been transferred, on the initiative of the Soviet Union, to the joint Comecon "Interkosmos" programme. At the same time, French participation will be largely reduced to the first stage of the mission, the fly-by of Venus.

The Halley probe is now to carry apparatus from Comecon countries, although their participation seems to be confined to the provision of equipment rather than the design of any individual experiments. Some countries may be assigned an individual responsibility for various sections of the mission — Hungary, for example, has been entrusted with developing the television system for monitoring the comet and transmitting the pictures to Earth.

The announcement of the Interkosmos trip to Halley's comet came at last month's meeting in Erevan of the Comecon permanent working group on space physics, only a few days after the successful landing of the Romanian cosmonaut, Dumitru Prunariu, brought to an end the programme of Comecon participation in missions to the Salyut space stations. Academician V.A. Kotelnikov, chairman of Interkosmos, confirmed that no further joint manned flights have been planned, although several unmanned joint missions are scheduled.

Vera Rich

tracking satellite would be able to cover the shuttle only in about half of its orbit and one third of Spacelab data could be lost. Extra recording equipment would be carried on board, but it is unlikely that the very high rate of data (100–150 megabits per second) from the microwave facility for Earth observations could be recorded. Spacelab users and ESA await NASA's final decision.

Meanwhile, NASA has been revising its schedule for shuttle flights up to 1985. Surprisingly, after the almost complete success of the first shuttle flight, delays are in store. Seven of the 44 flights scheduled before 1985 are to be postponed until the second half of the decade or even cancelled. Most other payloads are to be delayed. The reason for the delays, ranging from a few days to more than a year, are cuts in the Reagan budget, a more accurate estimate of the turn-round time and, most recently, delays in the manufacture of lightweight external fuel tanks. NASA says that delay in delivery of the tanks is largely responsible for the postponement of seven flights beyond 1985.

Three of the seven flights will effectively be cancelled. One will be saved by launching the two Galileo probes to Jupiter, more than 400 days later than originally planned, on a single shuttle flight. Another has been saved by the cancellation of the Venus-orbiting probe.

And the third may be saved from the International Solar-Polar Mission if NASA fails to win approval for even a modified solar-polar spacecraft. (The ESA solar-polar spacecraft is included in the shuttle manifesto but with a launch date of May 1986, more than a year later than originally planned.)

Foreign and commercial users affected by the latest shuttle delays are being given until the end of this month to decide whether they would like to use Thor Delta launchers instead.

Judy Redfearn

European innovation

Brokers in demand

Luxembourg

The role of the European Commission in creating the right environment for innovation came up again at a symposium — sponsored by the Commission — in Luxembourg last week. The three-day conference was inspired by the resolution of EEC science ministers at the end of 1979 and follows on from a similar symposium on banking and innovation held last year.

Despite the ritual homage paid to the idea that small business will make good the job losses in large companies and industries, this year's symposium highlighted the gap between the researcher and the businessman. Participants repeatedly affirmed that basic and pure research in Europe too often fail to lead to commercial application.

The implication is that the dynamism of the small American business is not so easily copied in Europe. Not only is there a lack of venture capital but linguistic differences, the small size of national markets and technical barriers are further deterrents. In addition, the sheer mass and variety of the channels through which research results are made available bewilder rather than inspire industry. So this symposium edged round to the conclusion that what is required is a broker or intermediary who could select exploitable research and help to develop it commercially.

Some effective middlemen exist. In the United Kingdom, the National Research Development Corporation was set up in 1946. The Netherlands has the Eindhoven University of Technology, which has helped 200 small and medium-sized businesses in its first year of operation. In Austria the Innovationsgesellschaft is a venture bank specializing in helping inventors to launch new products. Belgium has several industrial research parks at the Catholic University of Louvain-La-Neuve and Vrije Universiteit Brussels.

The consensus of opinion at the symposium, however, was that these innovations are inadequate. One suggestion was for a European Technology Transfer Centre, another for a European Investment Bank specializing in research. Others proposed more bulletins on the lines

of the French *Lettre des Sciences et Techniques* which selects and condenses reports of new discoveries. Many participants argued that the mentality of scientists in the field needs changing. They should publish information less for their peers than for the end user. Or they should themselves become entrepreneurs. The Commission will now consider how best to create the environment for the scientist/entrepreneur and for a European context for the scientific information broker.

Coincidentally, the European Commission's proposals for EEC programmes in the field of new technologies — microelectronics and telematics — are to be reexamined with greater interest, said Mr James Prior, the UK Secretary of State for Employment, after an unprecedented gathering of the EEC's finance and employment ministers in Luxembourg on 11 June with the aim of finding a solution to the growing unemployment problem. Sir Geoffrey Howe, the British Chancellor of the Exchequer, has also stated that the Commission is to bring forward new proposals to stimulate the expansion of small and medium-sized businesses.

If Mr Prior's words at the "Jumbo Council" are to be taken seriously, the Commission's work in the area of scientific communication is likely to be given a shot in the arm. The ministers agreed that there must be a greater coordination of the EEC's approach to new technologies.

Jasper Becker

Soviet scientists

Another trial

Dr Viktor Brailovskii, the Moscow cyberneticist faces trial this week on a charge of "disseminating fabrications . . . which defame the Soviet political and social system". Dr Brailovskii has refused a defence counsel, maintaining that nothing in his conduct over the past few years needs to be defended in court.

Dr Brailovskii was dismissed from his lectureship at the Moscow Radiotechnical Institute in 1972 when he applied, with his wife and son, to emigrate to Israel. In 1973, he and his wife Irina joined the "Sunday seminar on Collective Phenomena" organized by their friends, Dr Aleksandr Voronel, for "refusnik" scientists who wished to keep up some sort of intellectual life during the waiting period between applying for a visa (and subsequently losing their jobs) and actually being allowed to emigrate. From then on they were both frequently subjected to police harassment, which intensified when Dr Brailovskii became organizer of the seminar following the departure of Voronel and his successor Dr Mark Azbel' to Israel.

In 1976, Dr Brailovskii was given permission to emigrate but his wife was refused on the grounds that she had access to secret information. (Her former

immediate supervisor was prepared to deny this verbally, but not to commit himself in writing). On refusing to leave without his wife, Viktor Brailovskii had his permission to emigrate withdrawn in 1977.

In 1978, after a search of their apartment, his mathematical papers were confiscated and he was warned to stop lecturing on mathematics. In April 1980, the fourth in a series of conferences was to take place, basically an extension of the regular seminars to which foreign colleagues and sympathizers were invited. Before the 1980 conference Dr Brailovskii was arrested, questioned, and threatened with imprisonment if the meeting went ahead. Nevertheless it took place on schedule.

On 13 November 1980 Dr Brailovskii was arrested, and for several Sundays the police prevented anyone from going to his apartment. Dr Irina Brailovskaya, however, soon had the seminars going again, at first on Saturday evenings, and then, when the pressures lessened, at the original time on Sunday mornings. Dr Brailovskii, meanwhile, was held in Butyrki prison for more than 6 months (the official time for pre-trial detention is not more than three months, unless a special extension is sought).

Optical astronomy

Dutch sign on

Research councils in the Netherlands and the United Kingdom will sign an agreement today (18 June) which gives Netherlands astronomers 20 per cent of Britain's observing time at the Roque de los Muchachos Observatory (formerly called the Northern Hemisphere Observatory) in the Canary Islands.

In exchange for time on Britain's four planned telescopes — the 4.2-m William Herschel, 2.5-m Isaac Newton, a 1-m photometric telescope, and a 15-m millimetre-wave dish — the Dutch will provide 20 per cent of British costs and manpower in equipping and running the observatory.

The UK Science Research Council has estimated that the telescopes would cost £25 million in October 1979 prices, and that the running costs of the site will be of the order of £1 million a year.

The Netherlands has been hesitating over its contribution because of the sums involved, and also because it represents a certain shift of resources in favour of optical astronomy. The Dutch natural science research council (SWO) has previously done much of its optical astronomy through its membership of the European Southern Observatory, which has a fixed subscription; and Dutch radio-astronomers have voiced fears that the additional international commitment to the Roque will cut into their own budgets.

However, Dutch astronomers now agree that this level of commitment to Northern Hemisphere optical astronomy is

necessary. The world-famous Westerbork radio telescope is, after all, in the Northern Hemisphere, and more direct links with optical observations would be useful.

The observatory they thus become part of was set up by international agreement between Spain, the United Kingdom, Sweden, and Denmark, and though the Dutch enter by "the back door" — the new agreement being between research councils and not nations — such deals were allowed for in the original treaty. **Robert Walgate**

Trademarks and patents

More authorities

Brussels

Members of the European Federation of Pharmaceutical Industries' Associations have expressed anger at the European Commission's proposals to create a Community trademark office. The drug industry, which accounts for 65 per cent of all trademarks registered, feels threatened by what it sees as an attempt to ride roughshod over its interests.

Last November the Commission brought out a directive aimed at harmonizing the member states' trademark laws, with proposals for the creation of a Community trademark and an EEC trademark office.

Trademarks are one way of ensuring protection against imitators of new medicines and guaranteeing that drugs are properly used. The pharmaceutical industry's federation, meeting in Brussels in May, concluded that placing the registration of trademarks in the hands of the Commission could lead only to confusion, as national authorities will operate concurrently with the Commission. The federation would prefer to wait until national rules are more in line with each other before setting up a European trademark office.

There is much confusion in Europe not only over trademarks but over patents and the protection of intellectual property in general. This is an economic as well as a legal problem, for representatives of industry continually stress that inadequate protection of new products puts the profitability of new research at risk.

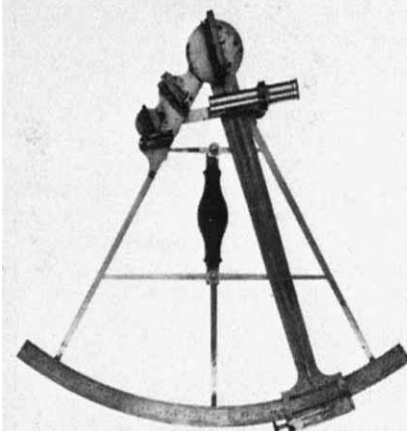
Although the European Convention on patents was signed eight years ago by ten European countries, three (Ireland, Denmark and Greece) have yet to ratify it. Italy has been particularly heavily criticized. Although it has ratified the convention, a new bill on drug patenting is being debated in the Italian parliament which appears to conflict with the convention. It is claimed that the Italian Senate has amended the bill to give protection to Italian firms manufacturing drugs which elsewhere would be protected by patent law. And in Yugoslavia manufacturers are said to be exporting low price drugs using patents provided by European countries solely for use in the domestic market. **Jasper Becker**

Cook's aid surfaces

A 15-inch radius brass sextant which is thought to have been taken on the last of Captain James Cook's historic voyages of exploration (1776–80), was sold at Sotheby's on 11 June for £11,000; it was made by Jesse Ramsden around 1775.

Ramsden was born in Halifax, Yorkshire, in 1735. His mathematical studies were interrupted by his father apprenticing him to a cloth-worker in Halifax; he then worked until about 1758 in a cloth-warehouse in London. It was here that he became apprenticed to a relatively unknown instrument maker named Barton (or Burton) in Denmark Court, Strand.

In 1763 Ramsden set up business on his own account in Haymarket and in 1775 he moved to Piccadilly, London, where he remained until his death in 1800. From the outset of his business life, his skill as an engraver and divider attracted the attention of leading London makers such as Nairne, Sisson, Adams and Dollond.



His design and construction of one machine revolutionized the dividing of the scales of sextants and other instruments. The Board of Longitude (set up by the Admiralty in 1714) awarded Ramsden £615 on the understanding that he would undertake to divide octant and sextant scales for other makers.

There is little doubt that it was directly due to the inventive genius and practical skills of Ramsden that the sextant was improved to the point that an accuracy of 10 or 20 seconds rather than 1 minute became the order of the day. In addition, the finely engraved scales made possible by the use of the dividing machine permitted sextants to be made smaller than previously without loss of accuracy.

Ramsden's improvements to the sextant eased life for the sailor, for latitude could now be determined with an unprecedented degree of accuracy. Even longitude, which had seemed to present insurmountable problems, was in the process of being solved, the solution being the provision of a suitable timekeeper, an accurate sextant and a nautical almanac. **Arthur Frank**

CORRESPONDENCE

Role of US Academy

SIR — Your editorial of 30 April (page 723) commenting on the recent history of the National Academy of Sciences under the leadership of Philip Handler was generous indeed. Nevertheless, it calls for some supplementation, lest your readers carry away a seriously deficient understanding of this institution's present concerns. The editorial urges the Academy to undertake study of a range of problems, both old and new, when — in fact — such studies are long underway. Some have already been completed.

"The education of students in schools and colleges?" or "The problem of university scientists, with or without tenure, engaged in the hazardous gamble of the pursuit of scholarship?" Our Commission on Human Resources has provided major studies of both these areas to the Federal Government within the past year.

"The relationships that should obtain between universities and the federal government with respect to the support and conduct of research? The impact on education of the new strict interpretation of rules prohibiting the export of military material?" The appointment of blue-ribbon committees to examine all aspects of these issues and to report to the Council of the Academy was announced several months ago.

In addition, a committee is being formed to look into the propriety of the arrangements academics make with universities to share in the profits of extramural commercial enterprises. With regard to new regulations from the Office of Management and Budget on time-accounting, the officers of the Academy and the membership at large have taken several steps to present their views to appropriate government officials, both publicly and privately.

Most baffling to us is the final question, "whether . . . to protest at Sakharov's exile and, if so, how?". Although many individuals and institutions have registered strong protests against the official harassment of Andrei Sakharov, surely the National Academy was among the first to express its grave concern over the treatment of its distinguished foreign associate — first privately and, since 1973, most publicly. Following Sakharov's internal exile, the Academy Council cancelled a symposium planned jointly with the Soviet Academy and deferred all other such symposia for a period of six months. With Sakharov's status unchanged, the Academy Council reaffirmed that position in August 1980 and just over a month ago.

Permit me one further observation. No one could disagree with your observation that the Academy should arrange to "shuffle off commissions for studies which are either pointless or untimely". Indeed, to reduce that likelihood, President Handler reorganized the National Research Council early in his first term to insure that all new activities of the Research Council had first been evaluated and authorized by appropriate bodies composed primarily of Academy members. Even so, there will always be disagreement over the relative significance of new studies. What is significant to A may appear quite trivial to B. For example, if your editorial meant to cite as

pointless and untimely the study of chlorofluoromethanes and the ozone layer, and of the health effects of dietary cholesterol and low-level ionizing radiation, Dr Handler, I know, would strongly disagree.

Even more fundamental is our belief that in a pluralistic society an institution is not likely to be permitted to answer only important questions.

HOWARD J. LEWIS

Director, Office of Information,
National Academy of Sciences,
Washington, D.C., USA

Questionable results

SIR — In answer to your editorial question "Who, not Congress, should police fraud?"¹, a recent proposal by Donald Goodwin² deserves consideration. He suggested that a Committee for Replication composed of scientists should be established. Each year the committee would select one or two findings they deemed important and surprising enough to warrant a special type of replication: replication at the laboratory from which the original finding came, but with an independent investigator present as a researcher, witness, and reporter. The independent investigator would be a scientist chosen by the committee from within the same field of research but from a different laboratory.

Goodwin's proposal has several merits:

(1) The process is directed towards establishing the reliability of dramatic new findings — the primary concern for scientists — and not specifically towards the discovery of fraud. Incorrect findings may occur in print for a variety of reasons, probably the rarest of which is fraud; the procedure would help expose incorrect findings regardless of cause.

(2) The process is more accurate in some ways than replication at other laboratories: for example, there may be critical differences between the procedures used at different laboratories that are not clear from the published accounts. Furthermore, in many fields it is rare that a pure replication is attempted; usually there are modifications, "improvements", introduced by the other laboratories. There has been generally only a low reward for conducting replications and perhaps as a consequence, many experiments were not duplicated³. The ratio of reward to effort for the independent investigator would probably be much higher with Goodwin's procedure, which also would assure the replication of the important experiments the committee selected.

(3) As mentioned in your editorial¹ "the scales are already too powerfully weighted against the heterodox". This procedure would not add to the weight against surprising findings and might help to correct the balance. Confirmation of a dramatic result by on-site replication would increase its acceptability and help to communicate the finding to a wider audience. It would also usually produce faster replications than relying on other laboratories.

(4) The process avoids the stigma of the witch-hunt. It would be an honour to have one's research chosen for on-site replication, a demonstration that the scientific community

recognizes the great importance of the results. It would provide an opportunity for an honest researcher to show the reliability of the results to sceptics. Nevertheless, the procedure would still act as a deterrent to fraud.

(5) Although the process would need to be funded, probably by a government agency, the actions would still be controlled by scientists. It would still be in the realm of peer review.

Such on-site replication is not, of course, a panacea, and there are several difficulties associated with it. From my own experiences as the independent investigator in the on-site replication attempt⁴ that initially stimulated Goodwin's proposal, I am quite aware of some of these difficulties, but I also can vouch for the overall feasibility of the procedure.

J.D. SINCLAIR

Research Laboratories of the State Alcohol
Monopoly (Alko), Helsinki, Finland

1. *Nature* **290**, 433-444 (1981).
2. Goodwin, D.W. at *XII Annual Nordic Meeting on Biological Alcohol Research*, April 21-23, 1981 Stockholm, Sweden.
3. Broad, W.J. *Science* **212**, 421 (1981).
4. Sinclair, J.D. at *XII Annual Nordic Meeting on Biological Alcohol Research*, April 21-23, 1981.

Film loop evolves

SIR — In his review of the Natural History Museum's new exhibition, *Origin of Species*, Professor Cox (*Nature* 4 June, p.373) refers to a "film loop" dealing with the status of the theory of evolution by natural selection. He regards this audio-visual programme as a failure and, apparently, misconceived.

In producing this audio-visual we were trying to avoid dogmatism in our presentation of the theory of evolution by natural selection. The intention was to reproduce the arguments set out in Chapter 12 of Colin Patterson's book, *Evolution* (British Museum (Natural History), 1978). Before the opening of the exhibition the Museum was aware that as a consequence of compression of the subject matter and transference to a new medium the audio-visual might give an impression other than that intended. Subsequent experience of the audio-visual as part of the exhibition has shown this misgiving to have been justified. A new version is being prepared.

R.S. MILES

Department of Public Services,
British Museum (Natural History),
London SW7, UK

Had enough too?

SIR — It is high time that you terminated the correspondence on the "philosophical" status of evolution. If some learned and pompous ass asserts that Darwin's theory is not "really" scientific, or that it is "socially embedded" or that holders of conflicting theories cannot "really" talk to one another about the same phenomena, then the necessary and wholly sufficient reply is one good old word that you probably wouldn't print.

M. HAMMERTON

Department of Experimental Psychology,
University of Oxford, UK

NEWS AND VIEWS

Gene transfer in plants

from Roy Davies

PLANT BREEDERS often need to introduce a single character from one genotype into another; a classic example is the introduction of a different gene or allele for disease resistance into an existing high-yielding cereal variety to combat an emergent virulent strain of a pathogen. The usual way of doing this is by recurrent backcrossing, but the generation times of higher plants makes this a slow and laborious procedure. In this issue of *Nature* (p.586), Jinks and his colleagues at the University of Birmingham describe an alternative and more rapid method of achieving this transfer of one or a few attribute(s) from one genotype to another.

Two diploid inbred lines (V27 and V12) of *Nicotiana rustica* were used in the experiments, with V27 being used as the maternal parent in the crosses and pollen of the V12 parent being exposed to doses of 0, 10, 15 or 20 krad of γ radiation. The two genotypes differ in several major genes; V27 plants are yellow, have a mop inflorescence and yellow ovaries, each of these characters being controlled by more than one recessive gene. V12 plants are green, have a non-mop inflorescence and black ovaries. In addition, the two lines are distinguishable in terms of several quantitative characters. For all the attributes listed, V12 has the dominant and V27 the recessive genes. As expected, the first generation plants derived from the crosses involving irradiated pollen (M_1 generation) were more variable than the F_1 . Plants of the M_2 generation, while still showing some variation, resembled the maternal parent, V27, more than did the F_2 , both for those characters determined by major genes and for the continuously variable characters. This resemblance increased with dose administered to the pollen. M_2 's resemblance to V27 is contrary to expectation since the pollen parent V12 contained the dominant alleles for the characters studied.

Up to this point the data could be interpreted in terms of a radiation-induced loss or mutation of the dominant alleles of the V12 parent, thus making the offspring

more like the parent having the recessive genes, V27. However, it is the other features of the data that make the experimental results of particular interest. Of 97 M_2 plants derived from the 20 krad treatment, 16 had the yellow plant colour of V27 and 10 were also mophead — a combination not found in the 100 F_2 plants examined. (There are several genes involved in the regulation of each of these characters.) Of these 10 plants resembling the maternal parent, that is, which were yellow mop, 6 had black ovaries — a paternally derived attribute. In terms of the three quantitative characters scored, all 16 yellow plants were very similar to the maternal parent. From previous experiments (Virk *et al. Heredity* 39; 287, 1977) the authors are confident that true-breeding yellow, mop, black-ovaryed genotypes can be generated in later generations. Thus from a small M_2 generation, plants could be isolated which were similar to the maternal parent, V27, for the two characters controlled by major genes, and for the three quantitative characters scored, but incorporating a single attribute from the paternal parent. Again an M_2 plant similar in all respects to V27 apart from one V12-like quantitative character was identified.

Comparable observations have been reported by Pandey in other species of *Nicotiana* (*Nature* 256; 310, 1975; *Heredity* 45; 15, 1980). He succeeded in making analogous transfers of particular characters both within and between species after exposure of pollen to ionizing radiations, albeit using much higher doses even than those used by Jinks and his colleagues.

These two sets of results are of considerable interest both because of their practical implication for plant breeding and for the mechanism involved. There is now a real incentive for plant breeders to

examine the extent to which this technique can be used for species other than those belonging to the genus *Nicotiana*, not only for the more rapid transfer of characters in intra-specific crosses, but also for the transfer of characters in those inter-specific or inter-generic crosses where, due to the incompatibility of the two genomes involved, the embryo aborts early in development. The introduction of only one or a few characters from another species or genus, rather than a whole alien genome, may well be tolerated, and allow the normal development of an embryo. There is yet another avenue to be explored; the production of somatic hybrids between species and genera of plants by the fusion of protoplasts is a rapidly developing technology. The hybrids themselves are unlikely to be directly useful as new crop plants, but by simulating the experimental approach used by Jinks and Pandey, and irradiating one of the parental protoplasts before fusion with an unirradiated 'recipient', specific desirable attributes could be transferred. A recent paper by Dudits *et al. (Molec. gen. Genet.* 179; 283, 1980) illustrates a first step in this direction. X ray-irradiated protoplasts of parsley ($2n=22$) were fused with unirradiated carrot protoplasts ($2n=18$) and among the regenerated plants selected were some which had 19 chromosomes; these had some genetic markers from the irradiated donor.

What are the mechanisms involved? Pandey has shown that his 'transformed' plants are diploid and show no gross chromosomal abnormality; furthermore, where the two parents differed in chromosome number the 'transformed' offspring had the maternal parent's chromosome number. In the present *N. rustica* experiments, cytological studies have not been undertaken but the fertility and morphology of the plants indicates that they are neither haploid nor grossly aneuploid. It appears that irradiation of the pollen fragments the chromosomes, but pollen tube growth can still occur even after the high radiation doses. The

Roy Davies is Professor of Applied Genetics at the University of East Anglia, Norwich.

fragments are transferred to the egg and parthenogenetic development of the egg occurs. The chromosome number of the egg or early embryo then doubles and finally some of the fragments are incorporated into, or associated with, the maternal genome. A crucial point to be investigated is whether the chromosome doubling and parthenogenetic development that occur so readily in these *Nicotiana* species, can occur as readily in other genera.

The size and nature of the chromosome fragments from the pollen parent, and the mechanism of their integration or association is of great interest. Genetical studies of linked markers and of segregation patterns in later generations will help to determine the size and location of the associated or integrated fragments, and physical characterization of particular

DNA fragments will be feasible. A general assumption of radiation cytogenetics is that broken ends of chromosomes rejoin only with other broken ends; since the maternal parent is unirradiated such ends will not be available for rejoining with the fragments produced in the pollen. How else might integration occur, if this is indeed what happens? Fragments could be incorporated during replication, or during the exchange events that ultimately become recognized as sister chromatid exchanges; if this is so, then the average size of the fragments would be expected to be small. Transposable elements of the kind first recognized in McClintock's work on maize, and physically characterized, for example, in *Copia* in *Drosophila*, must have a chromosomal organization which allows integration into, and excision from, eukaryotic chromosomes; but it seems

unlikely that random fragments induced by radiation would have a similar organization. One of many puzzles posed by the *Nicotiana* results is the frequency with which transfer of attributes occurs; 6 out of 97 M_2 progeny generated after a 20 krad dose of γ radiation to the pollen closely resembled the maternal parent yet had the black ovary character from the paternal parent. It remains to be seen whether this is stable in later generations and whether black ovary is typical in its rate of transfer; if it is, then this represents a high rate of integration or association. Clearly there are a large number of issues to be resolved but these investigations by Pandey and by Jinks and his colleagues should now provide an incentive for a great number of workers to re-examine the basis and practical application of the phenomenon of gene transfer. □

Cell sorting, cell differentiation and cyclic AMP

from Adrian Tsang

IN the slime mold *Dictyostelium discoideum*, cyclic AMP is now believed to be directly involved in the processes of cell sorting and differentiation, as well as acting as a chemotactic agent in mediating aggregation. This is one of the conclusions that emerged from a recent meeting* on the development of cellular slime molds.

Amoebae of *D. discoideum* multiply as isolated cells. The developmental programme is set in motion at the onset of starvation, and after about six hours the amoebae aggregate in response to cyclic AMP signals. The multicellular structures, called slugs, which are formed after aggregation are differentiated along the anterior-posterior axis. The anterior, prestalk region gives rise to the stalk cells of the terminally differentiated fruiting body, while the cells of the posterior region are the precursors of the mature spores. This formation of a patterned arrangement of new cell types has obvious analogues in the development of higher eukaryotes.

It has been known for some years that prestalk and prespore cells can sort out, but it is not entirely clear whether sorting out precedes differentiation or is a consequence of it (Forman and Garrod *J. Embryol. exp. Morph.* **40**; 229, 1977). The answer to this question is important because it will indicate whether or not the position a cell occupies in the aggregate has an effect on its differentiation. Making use of the observation that amoebae grown in a non-glucose medium, when mixed with amoebae grown in a glucose-containing medium, differentiate preferentially into

prestalk cells (Leach *et al. J. Embryol. exp. Morph.* **29**; 647, 1973), Takeuchi (Kyoto University) examined the relationship between sorting out and differentiation. He mixed ^3H -labelled, non-glucose-grown amoebae with glucose-grown amoebae and allowed them to form agglomerates in submerged conditions. At various times the agglomerates were sectioned and stained with anti-spore serum to locate the prespore cells, and the same sections were autoradiographed to locate non-glucose-grown cells. Takeuchi found that the prespore cells began to differentiate long before the labelled and unlabelled cells were sorted out, suggesting that cell sorting is not a prerequisite for differentiation. This thesis is supported by a separate experiment performed by Okamoto (Kyoto University) in which dissociated aggregative cells were shaken at high speed. Under these conditions, small clumps of cells were formed and the prespore cells in the clumps were found to be randomly distributed. As no cell sorting can occur in these small clumps of cells, the above data indicate that *D. discoideum* cells can differentiate in the absence of cell sorting. In other words, whether a cell differentiates into a prestalk or a prespore cell may not depend on the position it occupies in the multicellular aggregate.

Cellular sorting out has often been attributed to differential adhesion. That this may not be the whole story in *D. discoideum* was demonstrated by Sternfeld (Albert Einstein College) and Takeuchi. Sternfeld found that when dissociated slug cells were allowed to form clumps in suspension, a new pattern was formed with the prestalk cells in the centre of the clump. However, if the culture medium contained

cyclic AMP, the prestalk cells were found on the periphery. Moreover, Takeuchi showed that prestalk cells possess higher motive force than do prespore cells (Inouye and Takeuchi *J. Cell Sci.* **41**; 53, 1980). Although these experiments do not rule out the possible participation of differential adhesion in sorting out, they do strongly support the earlier suggestion of Matsukama and Durston (*J. Embryol. exp. Morph.* **50**; 243, 1979) that cell sorting is guided by differential chemotaxis to cyclic AMP.

Evidence accumulated over the past decade indicates that cyclic AMP participates in the differentiation of stalk and spore cells. Both Sussman (University of Pittsburgh) and Rutherford (Virginia Polytechnic Institute) showed that the potential of establishing a cyclic AMP gradient exists in the multicellular aggregate. Blumberg (MIT) and Okamoto provided additional evidence that the accumulation and the maintenance of post-aggregative gene products require cyclic AMP (Town and Gross *Devl Biol.* **63**; 412, 1978), and there is direct evidence that cyclic AMP affects the expression of developmental gene products by altering the rates of transcription of specific genes (Williams *et al. Proc. natn. Acad. Sci. U.S.A.* **77**; 7171, 1980). What remains obscure is the role of cyclic AMP in cell differentiation. Is it a general effector of gene expression, a morphogen which channels the cells into specific differentiative pathways, a maturation factor for stalk and spore cells or a combination of these? □

Adrian Tsang is at the Imperial Cancer Research Fund Mill Hill Laboratories, London, UK.

*The US-Japanese conference on the development of cellular slime molds held in San Diego, California in February was organized by I. Takeuchi (Kyoto University) and D. Francis (University of Delaware).

Scratches on the Palaeolithic record

from C.S. Gamble

THE study of our earliest ancestors often reveals a preoccupation with thoughts of food. In particular, the importance of meat in the diet of early hominids and the means by which it was obtained and shared has led to conflicting accounts about the cooperative¹ and aggressive² origins of the human species.

The difficulty of deciding between these alternative pictures of the past suggests that occasionally our understanding of early man owes more to our contemporary imaginations than to our ability to interpret the archaeological record. Current research is now tackling on an interdisciplinary front the complex issues of how this record is formed and how we can infer patterns of past behaviour from such contemporary observations.

An example of this approach is to be found in the continuing investigation of the early man sites at Olduvai Gorge, Tanzania. The studies reported in this issue of *Nature* by Bunn (p.574) and by Potts and Shipman (p.577) show how the well known data from these sites³ can still provide surprises when squeezed a little harder. In her landmark report, Leakey described many of the excavated sites at Olduvai as 'living floors'. This term implies that the bones and chipped stones found there were in some way causally related and that these observations could be used to reconstruct patterns of hominid behaviour that involved tool use, and its relationship to dietary behaviour.

But what if these and other East African sites were merely hydraulic jumbles⁴ where the association of stone tools with animal bones was the result of natural agencies, such as river transport, rather than hominid activity? Moreover, even if these natural agencies could be discounted, how could we distinguish between the bones that resulted from carnivore as opposed to human food refuse? The new observations by Potts, Shipman and Bunn provide good evidence that at least some of the bones and chipped stone tools were found together as a result of hominid rather than hydraulic activity. They have identified cut marks on the surface of herbivore long bones, which can be shown by use of a scanning electron microscope to correspond with those produced under experimental conditions by stone tools. The experimental work classifies various scratches as cut, slicing, scraping and chop marks, which can be distinguished from carnivore chewing, gnawing and bone crushing and from the distinctive gnawing patterns of rodents such as the porcupine. Scratches left on

bones by carnivore claw marks have not yet been examined.

Most of these cut marks have been found superimposed over carnivore scratches, but in at least one case this relationship is reversed, indicating that at least once a hominid got to the bone before the lions, hyenas and other carnivores of 1.8 million years ago. At the famous Zinj site at Olduvai, Bunn found that, from among a total of 3,500 animal bone fragments, about 300 bore cut marks while a further 400 were chewed by carnivores. Additional evidence comes from the recent excavations at Site 50, Koobi Fora in Kenya⁵. Here it was possible to fit together bone fragments, several of which also bore cut marks. The fragmentation of these long bones appeared in some cases to be related to the cracking of bones for marrow. Cut marks have also been found on bones which come from excavated sites where stone tools are not present.

These findings undoubtedly add a further dimension to our understanding of the archaeological record of early man. First, the behavioural association between stone tools and animal bones seems substantiated for at least part of these excavated assemblages. Second, the closer examination of carnivore gnawing on bones points to the composite nature of these sites and the fact that archaeologists must continue to investigate ways of distinguishing between the archaeological and palaeontological records⁶. These observations of cut marks are also important because of the difficulties of analysing traces of wear on stone tools. The recognition of distinctive polishes on the edges of flint artefacts which result from cutting and working meat, hides, wood, bone and plant material⁷ has opened up many research possibilities, but this is not yet possible for the lava, quartz and quartzite tools that make up the majority of tools from Olduvai and Koobi Fora. The proposal that these tools were used for stripping meat from carcasses and as hammers for cracking bones for marrow rests on the observation of cut marks and bone fracture patterns. Supporting evidence for the dietary practices of early man may also come from scanning electron microscope studies of fossil hominid teeth. Studies by Walker⁸ of baboons and hyraxes has shown that different diets lead to different attritional patterns on tooth surfaces and these investigations are now being applied to specimens of fossil man⁹.

What do these observations contribute to our understanding of the behaviour of early man? Although they show that meat and marrow were processed and consumed, they do not directly answer the question of how, through either hunting or

scavenging, this food was obtained.

This is a problem of inference. How can archaeologists reconstruct the dynamics of past hominid behaviour when they can only work on the inert stone tools and bones that they dig up? This problem is further complicated in the study of early man since there are no contemporary examples providing suitable behavioural models to help interpret the excavated archaeological patterns. We can argue that the association of stone tools and animal bones demonstrates that meat was eaten and suggest that it was acquired by hunting or scavenging; but although it is a reasonable explanation for such archaeological observations, we are still mostly guessing at what went on in the past.

L.R. Binford¹⁰ considers that one of the commonest residues found on archaeological sites, the structure of animal skeletons, provides a constant against which some aspects of hominid and carnivore behaviour can be measured. Observations of skeletal disarticulation and bone destruction by carnivores and contemporary hunter-gatherers¹¹ reveal very different patterns that can be related to uni-formitarian principles of the mechanics of feeding. Using the work of A.P. Hill on the taphonomy of East African large ungulates and his own observations, Binford shows that it is possible to distinguish between residual assemblages of animal bones that have been primarily utilized by man and by carnivores. Moreover, he shows how these differences can be accounted for in terms of a predictable set of variable behaviour. In other words, it is possible to take a collection of inert observations from the archaeological record and infer the dynamic pattern of past behaviour that resulted in those residues.

Binford has applied this methodology to the 'living floors' from Olduvai Gorge. He concludes that the anatomical parts of the large animals that were found in these sites fit the model of carnivore feeding and that we are looking at ravaged carcasses. Not all of the patterning from these animal bone collections could, however, be explained

C.S. Gamble is in the Department of Archaeology, University of Southampton.

1. Leakey, R. & Lewin, R. *People of the Lake* (Avon, New York, 1978).
2. Ardrey, R. *The Hunting Hypothesis* (Atheneum, New York, 1976).
3. Leakey, M. *Olduvai Gorge* Vol.3 (Cambridge University Press, 1971).
4. Binford, L.R. *J. anthrop. Res.* 33 (1977).
5. Bunn et al. *World Arch.* 12 (1980).
6. Behrensmeier, A.K. & Hill, A.P. (eds) *Fossils in the Making* (University of Chicago Press, 1980).
7. Keeley, L. *Experimental Determination of Stone Tool Uses* (University of Chicago Press, 1980).
8. Walker, A. *Fossils in the Making* (eds Behrensmeier, A.K. & Hill, A.P.) (University of Chicago Press, 1980).
9. Walker, A. *Phil. Trans. R.Soc. B292*, 57 (1981).
10. Binford, L.R. *Bones: Ancient Men and Modern Myths* (Academic, New York, 1981).
11. Burford, L.R. *Nunamiut Ethnoarchaeology* (Academic, New York, 1978).

by the carnivore model — some marrow bones which should have been present were missing and the differential pattern of bone destruction favoured additional agencies involved in splitting bones for marrow. This is the information about hominid behaviour. These hominids were scavenging the kills of large predators for bones of low food utility, scraping off the last scraps of meat and cracking them for marrow which the carnivores had ignored. Transport of the scavenged bones away from the kill site and to more protected locations also took place. Stone tools to

crack open these bones and to slice off the last pieces of meat gave man a particular advantage in competing for this niche that utilized the waste of the predators. The evidence for cut marks and the investigation of bone breakage patterns would seem to support this reconstruction. The strong inferential base of this behavioural reconstruction now allows us to measure the activities of early man against the evidence from the rich palaeontological record of which the archaeological evidence forms a small but significant part. □

positive signal of fractional charges in niobium spheres of radius 1/10 millimetre (LaRue *et al. Phys. Rev. Lett.* 15; 967, 1981). Of the 39 measurements on the 13 spheres tested, 14 values of fractional charge are observed on 5 different spheres. Furthermore, all 39 measurements give values of the charge consistent with zero or $\pm e/3$. (In these experiments, charges of $\pm 2e/3$ are indistinguishable from $\pm e/3$). Does this mean that free single quarks have been observed? If so their density in niobium is at least 1 quark per 10^{20} nucleons. Would this be in conflict with earlier null measurements?

There still exists the possibility that, since fractional charges have been seen in only one experimental apparatus, they are in some, not understood way an instrumental effect. For example, it is necessary to apply corrections of magnitude larger than e to the observed measurements in order to deduce the true values of the charge, and also some measurements are discarded due to apparatus instabilities during the course of a run. In order to eliminate the possibility that such effects could enable the experimenters (subconsciously) to adjust their final measured charges to be $\pm e/3$, future measurements by the Stanford group will have a random extra fractional charge inserted into their data, and whose magnitude is unknown to them. Only after they have decided whether to accept a measurement and have completed calculating their corrections will the random charge be subtracted. If they then still find values consistent with only zero and $\pm e/3$, the evidence that the fractional charges are genuine will be much more convincing.

The Stanford experiment's positive result is not necessarily inconsistent with the many previous null results because there are no other direct measurements on niobium to date. As far as searches in other materials are concerned, there is no particular reason to suppose that the quark/nucleon density is independent of material. However, Morpurgo points out that the Stanford spheres change their charge quite readily; this is interpreted as quarks dropping off the surface from time to time, so they should be found in other materials as well. Clearly it is desirable to repeat the search for fractional charges in niobium by as direct a technique as possible in another apparatus.

It is also possible that the Stanford fractional charges are nothing to do with the quarks that are responsible for building elementary particles, but correspond to something else altogether; indeed the Stanford paper makes no mention of the word 'quark'. If, on the other hand, free quarks have been seen, it would significantly modify current theoretical ideas about the way quarks interact with each other. Meanwhile the experimental situation needs clarification. The outcome is awaited with great interest. □

The discovery of free quarks?

from Louis Lyons

SINCE earliest times there has been speculation on the nature of the ultimate constituents of matter which are responsible for the enormous variety of objects that exist in the Universe. By the beginning of this century, it was believed that there were just two 'elementary particles' — the proton and the electron — out of which everything was constituted. This satisfactory state of affairs was gradually eroded with the discovery of more and more particles, so that today the list of so-called elementary particles contains several hundred entries. There thus exist more 'elementary particles' than chemical elements.

In 1964, Gell-Mann (*Phys. Lett.* 8; 214, 1964) and Zweig (CERN preprint TH412) independently suggested that the multitude of elementary particles were made out of simpler objects called quarks. At the time only three quarks were needed to explain the existence and many of the properties of all known particles, as well as the absence of certain conceivable types of particle which have not been found. Subsequently the quark model was also used to explain many of the features of scattering experiments involving the composite 'elementary particles' (Feynman *Photon-Hadron Interactions* Benjamin, New York, 1972).

The simplicity and success of the quark model immediately lead to a succession of experiments designed to detect free quarks. Despite the fact that all elementary particles have electric charges which are integers (when measured in units of e , the electron's charge), most versions of the quark model require quarks to possess fractional charges of $\pm e/3$ or $\pm 2e/3$. Free quarks should also possess masses different from the conventional elementary particles. Experimental observations of unusual charges and/or masses could be evidence for the existence of free quarks.

Despite the variety of tested materials (which have included the Sun, moon soil, meteorites, ocean sludge, mountain lava, oysters, cosmic rays and particles produced in high-energy interactions at accelerators), most experiments have provided no convincing evidence for the existence of free quarks. Thus Morpurgo and his collaborators in Genoa have used a modification of the original Millikan oil drop experiment to measure the charge on steel balls. They find no fractional charges and set a limit of less than 1 quark in 10^{21} nucleons.

Other experiments have set even lower upper bounds on quark concentrations, but such experiments have involved quark enrichment and transportation techniques that rely on assumed chemical and physical properties of quarks. A positive signal from such experiments would have been significant, but the negative results may simply mean that free quarks do not behave in the assumed way. Quark search experiments have recently been reviewed by Jones (*Rev. Mod. Phys.* 49; 717, 1977) and by Lyons (*Prog. Particle Nucl. Phys.* 7; 157, 1980).

The almost universal absence of an observed signal of free quarks leads to the concept of 'confinement': quarks are able to exist in pairs and in triplets inside elementary particles, but it is impossible (according to this hypothesis) to pull them apart and observe them singly. Perhaps a quark is like one end of a piece of string; no matter how hard one tries, it is impossible to isolate a single end. Vigorous attempts are being made to prove that confinement is a necessary consequence of Quantum Chromodynamics (QCD), a theory of the interactions of quarks with each other. Despite the lack of success to date, confinement is still held as a popular belief (Johnson *Scient. Am.* 241; 100, 1979).

Recently, however, a Stanford experiment, using a different variant of the Millikan experiment, has reported a

Louis Lyons is in the Nuclear Physics Laboratory, Oxford.

Raman: an eye for combustion studies

from D.A. Greenhalgh

Accurate measurements of the events taking place in the very hostile environments inside combustors such as furnaces, gas turbines and internal combustion engines has not been possible until very recently. Now, a new generation of optical measurement techniques based on the laser are offering real solutions to the problems of making remote, non-intrusive measurements.

By far the most successful technique for temperature and specific species measurements is laser Raman spectroscopy; in particular, Coherent Anti-Stokes Raman Scattering (CARS) which has been used for measurement inside internal combustion engines and jet combustors. Together with laser anemometry, which provides information on gas velocities and turbulence, these techniques could well revolutionize our knowledge of combustion processes over the next decade.

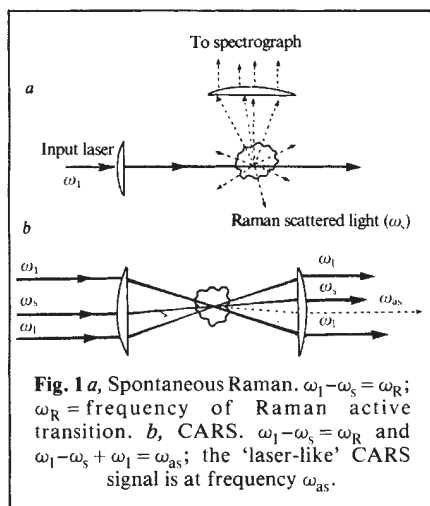


Fig. 1 *a*, Spontaneous Raman. $\omega_1 - \omega_s = \omega_R$; ω_R = frequency of Raman active transition. *b*, CARS. $\omega_1 - \omega_s = \omega_R$ and $\omega_1 - \omega_s + \omega_2 = \omega_{as}$; the 'laser-like' CARS signal is at frequency ω_{as} .

In order to appreciate the role of Raman spectroscopy in combustion, it is necessary to outline the basics of a practical combustor. The fuel is frequently a liquid hydrocarbon which is sprayed or directly injected into an air flow, and burns after a suitable mixture of vaporized fuel and air has formed. The process is often described as 'diffusion burning' and the chemical kinetics are fast compared with the mixing processes. This is, however, a rather simplified picture — even before actual combustion, the fuel molecules will be pyrolysing to form lighter molecules and soot, or its precursors.

Coal combustion is considerably more complex but nevertheless it can be modelled in a predominantly aerodynamic manner, concentrating on parameters such as air flow and temperature. To predict these parameters accurately a moderate amount of chemistry has to be included in

D.A. Greenhalgh is in the Engineering Sciences Division, AERE Harwell, UK

the models. Whereas flame chemistry can be studied through measurements of well controlled laboratory flames, the physics of diffusion flames often requires information from actual devices or close replicas. Although Raman spectroscopy is important in measurement of both physical and chemical properties, the role of CARS for measurements in practical combustion flows will be emphasized here.

Raman scattering involves inelastic scattering of light by molecules. In conventional Raman spectroscopy an intense beam of monochromatic light is focused at the desired measurement point within the sample gas and the inelastic scattered light, which is Stokes-shifted from the incident light, is subsequently analysed with a spectrograph or similar instrument. Such a system only collects light from the focus of a laser beam and thereby ensures very high spatial resolution. From the spectra of simple molecular species, such as diatomics, water and carbon dioxide, it is possible to deduce both temperature and concentration of the individual component molecules. Unfortunately, conventional Raman spectroscopy is limited to studies of premixed flames or to diffusion flames, such as hydrogen and air, where soot formation is impossible. This restriction is not set by the luminosity of sooty flames but because the fluorescence induced by the incident laser light swamps the weaker Raman-scattered light.

New Raman techniques, of which CARS is the most successful, developed from research in non-linear optics, particularly that of Taran and his co-workers². In conventional Raman the scattered light obeys the relationship $\omega_1 - \omega_{sj} = \omega_{Rj}$ where ω_1 is the frequency of the incident light and ω_{sj} is the *j*th component of the Stokes-shifted Raman signal corresponding to each ω_{Rj} , the frequency of the *j*th individual Raman active transition. To obtain a CARS spectrum, three laser frequencies are introduced into the medium such that the signal at frequency ω_{as} (as = anti-Stokes) obeys the relationship $\omega_{as} = \omega_1 + \omega_2 - \omega_s$. Usually ω_1 and ω_2 are the same laser and we have the degenerate case where $\omega_{as} = 2\omega_1 - \omega_s$. The latter process is represented diagrammatically in Fig.1 where a comparison is drawn with conventional or spontaneous Raman. If $\omega_1 - \omega_s = \omega_{Rj}$, then the process is resonant and a strong 'laser-like' emission is produced at ω_{as} . In common with conventional Raman, very high spatial resolution ($\sim 2 \text{ mm} \times 0.2 \text{ mm}$ diameter) is possible. In conventional Raman the incident laser excites the molecules of the medium which then spontaneously decay, releasing the Stokes-shifted signal at ω_s . Due to the spontaneous

nature of the process, the signal is scattered into 4π steradians. By comparison, CARS is a coherent process, where the laser fields at ω_1 and ω_s mix within the medium to induce a coherent signal at ω_{as} ; thus the signal has the optical properties of the input lasers and is a 'laser-like' beam. This simple facet of CARS is its main strength; a laser beam can be collected efficiently by optics with a very small aperture whereas laser-induced fluorescence or a conventional Raman signal cannot. This advantage has been realized for several years and was demonstrated in sooty flames by Beattie and co-workers⁴.

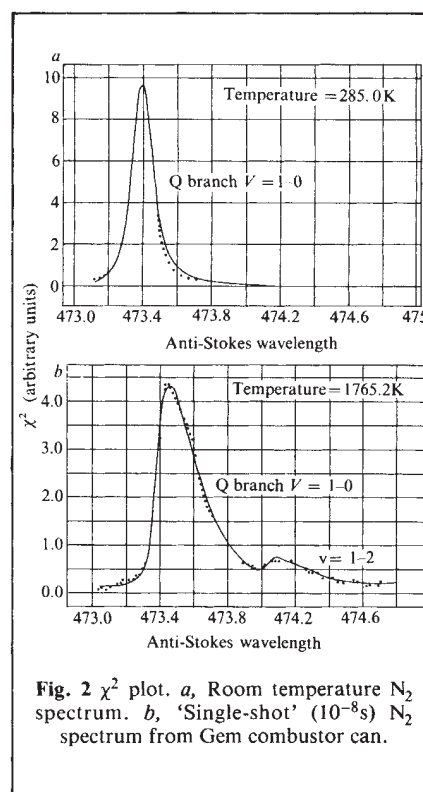


Fig. 2 χ^2 plot. *a*, Room temperature N_2 spectrum. *b*, 'Single-shot' (10^{-8} s) N_2 spectrum from Gem combustor can.

There is another fundamental difference between the two techniques: in conventional Raman a complete spectrum is generated by one monochromatic incident wavelength ω_1 , whereas only one frequency component, ω_{as} , of the CARS spectrum can be generated if ω_1 and ω_s are monochromatic. In order to generate a Raman spectrum by CARS, one laser, normally ω_s , is frequency tuned so that a spectrum of the Raman active transitions is produced. This approach has been used to demonstrate CARS in both a gas- and a liquid-fuelled internal combustion engine by Stenhouse *et al.*², and in a combustion tunnel burning kerosene fuel, by Taran⁶. Although both these experiments demonstrated the feasibility of temperature measurements from the Raman

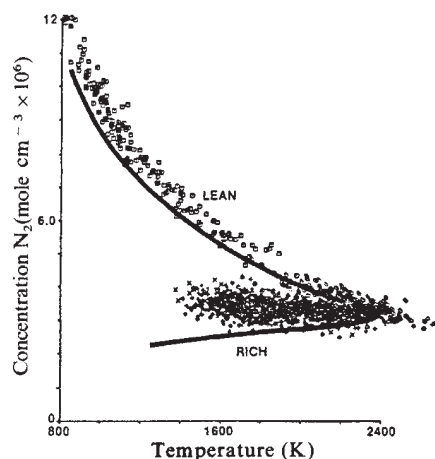


Fig. 3 Comparison of simultaneously acquired N_2 concentration and temperature data with a theoretical curve for a hydrogen air flame (taken from ref.9).

spectrum of nitrogen, Taran's work also demonstrated the possibility of CO_2 detection down to an estimated 1,000 p.p.m. and the measurement of CF_4 at around one per cent. The latter was injected into the combustion flow for short periods of time to infer residence times.

This approach of obtaining a spectrum by scanning the frequency of ω_s is satisfactory for demonstration experiments and laboratory gas samples, but it can lead to errors in determining mean values of

temperature and concentration in fluctuating or turbulent conditions. Because the Raman spectrum is a non-linear function of temperature (it is a function of the Boltzman distribution of population of energy states), averaging varying spectra does not lead to a true mean temperature. In order to freeze these fluctuations in time, short-pulsed lasers are used (pulse length $\sim 10^{-8}$ s) and the complete spectrum is obtained in one shot. In conventional Raman this is simple since all of the spectrum is generated simultaneously, but for CARS to achieve the same result it is necessary to use a 'broad-band' laser at ω_s so that a complete section of a Raman spectrum is generated at the CARS signal frequency ω_{as} .

This approach has been very successfully developed by Eckbreth and co-workers⁷ who have demonstrated single-shot broad-band CARS measurements in a variety of flames. In particular, they have measured temperature in a Pratt and Whitney JT-12 combustion can burning Jet 'A' fuel. Similar studies have recently been completed in this country on a Rolls-Royce 'Gem' combustion can fuelled with gas⁸. A typical 'single-shot' flame and a room temperature nitrogen spectra are shown in Fig.2, together with least-square fits of theoretical spectra. These spectra clearly show a large variation with temperature. In this study a sufficiently large number of 'single-shot' spectra from each geometric

point was obtained to allow an estimation of both the mean and the r.m.s. of the temperature. Statistical information of this type is used to validate the aerodynamic type of mathematical model discussed earlier.

The above examples clearly demonstrate the potential for CARS spectroscopy for physical and chemical measurements in practical combustion devices, and suggest that CARS might replace conventional Raman as the preferred technique for combustion analysis. Complete replacement is unlikely since CARS is both financially and technically more demanding than conventional Raman, and the latter remains well suited to combustion measurements where signal strength and fluorescence rejection are not paramount. Indeed, fundamental combustion studies have clearly illustrated the wide-ranging ability of conventional Raman. Of particular interest is the work of Lapp and co-workers⁹ where data have been collected simultaneously for temperature and nitrogen, hydrogen and water concentrations for each laser shot in a hydrogen air diffusion flame. Correlated data of this type can be very usefully compared with theoretical models. Such a comparison for a temperature-nitrogen concentration correlation is shown in Fig.3. Models of this type assume fast chemistry with the consequence that the combustion process is dominated by fuel air mixing. Although results show good



100 YEARS AGO THOUGHT-READING

THE public mind has of late been somewhat agitated by the doings of a Mr Bishop, who has come before the world of London society in a capacity no less startling than that of a professed reader of thought. He has not only taken by storm the general public and daily press, but also succeeded in convening an assembly of scientific men to witness his performance, while still more recently he has had the honour of exhibiting his powers before the Heir Apparent to the Crown. It seemed to Prof. Croom Robertson worth while to make a careful trial of Mr Bishop's powers, and he therefore invited Mr Francis Galton, Prof. E.R. Lankester, and myself to join him in an investigation.

First, Mr Bishop was taken out of the room by me to the hall down stairs, where I blind-folded him with a handkerchief; and, in order to do so securely, I thrust pieces of cotton-wool beneath the handkerchief below the eyes. While I was doing this, Mr Sidgwick was hiding a small object beneath one of the several rugs in the drawing-room; it having been previously arranged that he was to choose any object he liked for this purpose, and to conceal it in any part of the drawing-room which his fancy might select. I then led Mr Bishop up stairs, and handed him over to Mr Sidgwick. Mr Bishop then took the left

hand of Mr Sidgwick, placed it on his (Mr Bishop's) forehead, and requested him to think continuously of the place where the object was concealed. After standing motionless for about ten seconds Mr Bishop suddenly faced round, walked briskly with Mr Sidgwick in a direct line to the rug, stooped down, raised the corner of the rug, and picked up the object.

It was soon found that Mr Bishop succeeded much better with some of us than with others; so at a second meeting, in order to make a numerical comparison, he was requested to try two experiments with each of the four persons who were present. With Mr Galton, Prof. Robertson, and Prof. Lankester he failed utterly, while with myself he succeeded once perfectly and the second time approximately. For on the first occasion I concealed a pocket-matchbox upon the top of a book behind the leather lap of a book-shelf. After feeling along the rows of books for some time he drew out the one on which the matchbox was lying. In the second experiment I placed a visiting-card on the key-board of a grand piano and closed the cover. After going about the room in various directions for a considerable time he eventually localised the piano, and brought his finger to rest upon its upper surface about six inches from the place where the card was lying. It has also to be mentioned that in one of the experiments which he tried with Prof. Robertson the evening before, he was, after a good deal of feeling about, successful in localising a particular spot on an ordinary chair which Prof. Robertson had selected as the spot to be found. From this it will be seen that it made no difference whether a particular article or a particular spot was thought of; for

if the subject thought of was a certain square inch of surface upon any table, chair, or other object in the room, Mr Bishop, in his successful experiments, would place his finger upon that spot. Neither did it make any difference whether the article or place thought of was at a high or a low elevation. Thus, for instance, in one of the experiments I placed a small pencil-case high up in the chandelier of one of the drawing-rooms. While feeling over the surface of a table in the other drawing-room, and not far from the corresponding chandelier, Mr Bishop suddenly pointed at arm's length vertically to the ceiling. He remained motionless in this position for a few seconds, and then set off at a brisk pace in a straight line to the other drawing-room, until he came beneath the other chandelier.

From this brief summary it will be seen that on the whole his power of localising objects or places thought of by a person whose hand he clasps is unquestionably very striking. Of course the hypothesis which immediately suggests itself to explain the *modus operandi* is that Mr Bishop is guided by the indications unconsciously given through the muscles of his subject. He describes his own feelings as those of a dreamy abstraction or "reverie," and his finding a concealed object, etc., as due to an "impression borne in" upon him. But however this may be all our experiments have gone to show that Mr Bishop owes his success entirely to a process of interpreting, whether consciously or unconsciously, the indications involuntarily and unwittingly supplied to him by the muscles of his subjects.

from *Nature* 24, 23 June, 171, 1881.

agreement for the lean zones of combustion, there are some discrepancies in the theory for the rich combustion zones. However, this type of approach will be of considerable value to the development of combustion models.

All the previous examples are concerned with turbulent diffusion flames where fuel air mixing is the dominant feature of the combustion process. Many of the fundamental chemical processes of combustion have been deduced from studies of premixed lamina flames, where the combustion process is stable in space and time. Conventional Raman spectroscopy can be readily applied as long as the stoichiometry is arranged for an excess of air so that soot formation is largely suppressed; for example, Bechtel and Blint¹⁰ have examined the differences in flame composition in the main stream of a combusting flow and close to a wall. This type of study is of great interest since most practical combustors have cooled surfaces which undoubtedly affect both the physics and chemistry of the flow.

With regard to the wider applications of Raman techniques for gas-phase studies, consider the comparison between a gaseous chemical reactor and a combusting flow. In the former case the temperature is lower ($\sim 800\text{K}$ compared with $1,700\text{K}$) and the reactants are not necessarily hydrocarbon and oxygen. However, the flow is often turbulent and particle laden as in a (practical) combustor. Application of CARS to this area of research should help characterize chemical processes inside gas-phase chemical reactors.

Recent developments in Raman spectroscopy have already moved the technique to the forefront of combustion research, and there remains considerable scope for its development in, for example, the area of gaseous chemical reactions for both homogeneous and gas-phase measurements above catalytic surfaces. Conventional Raman reactions will continue to contribute fundamental data from controlled laboratory systems; but the new generation of Raman techniques, in particular CARS, will make possible the difficult task of measuring temporally and spatially precise temperatures and concentrations in practical combustors.

Amazonian question

from D.H. Tarling

THE SOUTH AMERICAN SHIELD covers a vast area of some 8,000,000 km², stretching from northern Venezuela to northern Argentina. To the west it is bounded by the Andes and to the east, it originally linked with Africa. However, the area is still largely unknown because of the difficulty of access and the general lack of exposure of basement rocks. The shield poses many questions of both local and international importance, for example the apparently unique persistence of Archaean-type tectonics well into Proterozoic times, which were recently discussed at the First Amazonian Symposium held at Puerto Ayacucho in Venezuela.

One of the fundamental economic problems is the apparent lack of the mineralization normally associated with Archaean terrains. The question is whether such mineralization never occurred, occurred and has not been located, or occurred but has since been eroded. In the case of the ultrabasics and greenstones in Brazil, Guyana and Surinam, the major problem appears to be that the techniques of exploration have not yet been sufficiently modified for local conditions — the great thickness (often 100 m) of laterite effectively prevents the use of most geophysical techniques while geochemical prospecting really requires a return to first principles, rather than the adoption of techniques developed for use in the medium latitudes of the Northern Hemisphere. G.W. Walrond (Guyana) identified seven distinct metallogenic provinces but the trace mineral associations that needed to be mapped may not be those used for the Northern Hemisphere where the degree and nature of hypergene weathering are so different. Nonetheless, it was clear that in this area, at least, the mineralization was present, but the current techniques were not really suited for its optimum location.

Such an assessment applies also to the use of remote sensing in such areas. In general, the vegetation in equatorial rain forests exists independently of the basement rock. It does not, therefore, show the usual stress features associated with, for example, copper impurities in the subsoil. Ecological provinces can be recognized from satellite imagery, but they are too large to be used directly in the search for minerals. Structurally controlled lineations can, however, be determined, particularly from radar scans of the region. The fact that the vegetation is self-dependent also

highlights the problem of the extent of forest burning in many parts of the Amazon Basin. Once destroyed, the soils and humus are rapidly removed and the area becomes sterile. In such remote areas, the problem is not one of legislation but one of enforcement.

Elsewhere Archaean mineralization may have existed, but is now eroded. Many of the richest mineral deposits associated with Archaean terrains are, in fact, sedimentary concentrations that have since metamorphosed. The original igneous sources of copper and gold, for example, are usually in the greenstones and ultrabasics, but these are marginally economic, and it is surficial concentration that caused the real enrichment of areas such as the Copperbelt of Central Africa. Such superficial deposits may have been removed from the Venezuelan part of the Archaean — the Imataca Province — which was possibly eroded to greater depths than similar greenstone terrains elsewhere. There is one simple test of this hypothesis. Pollack and Chapman (*Tectonophysics* 38; 279, 1977) demonstrated an age relationship between the rate of heat flow and the age of the last orogenic event in continental areas. Such a relationship requires a strong concentration of radiogenic heat-producing elements in the upper two or three kilometres of the continental crust. If the depth of erosion in northern Venezuela has been that deep, such radiogenic heat sources would have been removed and this area would now be characterized by even lower heat flow.

The most intriguing problem is the source of the diamonds that can be found extensively over the shield. These diamonds are mostly of industrial grade, but the few gemstones are of exceptional purity. It was not surprising, therefore, that Baptista and Parra's (Venezuela) report of diamond finds in the sediments of the Quebrada Grande river at 144,000 carats per km² caused considerable excitement. In South Africa, such concentrations would normally be very close to their kimberlite source. However, in South America these alluvial diamonds are not found associated with the other minerals usually found near kimberlite sources. Asbestos was found but is attributable to ultrabasic intrusives nearby; none of the chrome spinels associated with kimberlites was reported. It may well be asked why it is so essential to find the source as, in general, diamonds within kimberlite pipes are very expensive to extract. The reason is that knowledge of the source is required to assess areas in which other placed deposits are likely to occur. The search for

1. Raman, C.V. *Ind. J. Phys.* 2, 387 (1928).
2. Regnier, P.R. & Taran, J.-P.E. *Appl. Phys. Lett.* 23, 240 (1973).
3. Maker, P.D. & Terhune, R.W. *Phys. Rev.* 137, A 801 (1965).
4. Blattie, J.R., Black, I.D., Gilson, T.R. & Greenhalgh, D.A. *Coherent Anti-Stoke Raman Spectroscopy in Sooty Diffusion Flames* SPIE Pap. 158-09 (1978).
5. Stenhouse, I.A., Williams, D.R., Cole, J.B. & Swords, M.D. *Appl. Optics* 18, 3819 (1979).
6. Attal, B., Pealat, M. & Taran, J.-P.E. *CARS Diagnostics of Combustion*, presented at the AIAA 18th Aerospace Sciences Meeting, Pasadena, January 1980.
7. Eckbreth, A.C. *Comb. & Flame* 39, 133 (1980).
8. Greenhalgh, D.A., England, W.A. & Porter, F. to be published.
9. Lapp, M. & So, R.M.C. *The Study of Turbulent Diffusion Flames: Modelling Needs and Experimental Light Scattering Capabilities*, Proc. AGARD Conf. on Testing and Measurement Techniques in Heat Transfer and Combustion, Propulsion and Energetics Panel, Brussels, May 5-7, 19-1 (1980).
10. Bechtel, J.H. & Blint, R.J. *Appl. Phys. Lett.* 37, 576 (1980).

D.H. Tarling is a Reader in Palaeomagnetism, in the Department of Geophysics and Planetary Physics, Newcastle University, UK.

Opportunism of plant seeds

from Peter D. Moore

THE germination of many types of seed can be triggered by various external, physical stimuli, such as temperature, water availability and light. Logically it would seem that such response mechanisms have developed as a consequence of a selective advantage gained by the plants whose development has previously been constrained in this way. Dormancy thus prevents the plant from germinating and expending its reserve at an inopportune moment, that is, one in which its chances of long-term survival are small.

The germination strategies of plants are therefore closely linked to the ecological problems presented by their particular habitat. Many weed species occupying unpredictable habitats, in which frequent catastrophe is an inevitable component of life, rely strongly upon the maintenance of a reservoir of dormant seeds in the soil to ensure survival of the holocaust. In such species a simple technique for dormancy maintenance is the linkage of dormancy breakage with exposure to light. Species which grow in dense grassland habitats would be at a great disadvantage if dormancy were maintained at low light intensity, for at ground level they may

rarely experience anything else. The seeds of such species usually germinate in the dark. Species of woodland clearings depend on the death of a canopy tree or some other disturbance to open the leaf cover and provide an opportunity for growth. It might be supposed that such species would be at an advantage if their germination were finely tuned to the changes in spectral quality of light associated with a reduction in canopy. This is speculation, but it would make sound ecological and evolutionary sense.

Cresswell and Grime (see p.583 of this issue of *Nature*) point out an interesting correlation between the light demands of seeds and the environment within which they mature in the fruit. Seeds developing within tissues which are photosynthetic are shielded by these green layers and subsequently require a light stimulus for germination. Seeds which are exposed to full light before their final maturation and drying are capable of germinating without light.

Cresswell and Grime propose that species of unpredictable habitats which

gain advantage by maintaining a dormant population of seeds in the soil, retain chlorophyllous tissues around their maturing seeds, thus producing a light requirement. It is even possible that some species whose seeds do not become incorporated in the soil seed bank may retain a green cover over maturing seed and thus induce a light requirement; such species may require the full light of an opening in the canopy to trigger germination.

Obviously, light is just one of many factors in the environment which influence germination and it is not always the overriding one. However, the experiments of Cresswell and Grime do suggest that the morphology and phenology of the inflorescence, flower and fruit can have a direct influence upon the subsequent germination requirements of the seed. This leads to the interesting possibility that ecotypic development within certain species, providing adaptation to a variety of habitats, could relate to such a minor modification as the chlorophyll content of structures investing the seed. It will be of value to look for such modifications among those plants which are capable of survival in both stable and unstable environments.

Peter D. Moore is in the Department of Plant Sciences, University of London King's College.

kimberlites, however, persists. Elsewhere in the world the formation of kimberlites seems to require two factors — an Archaean-type upper mantle and a later addition of volatiles to it. In South America, Archaean areas are present, but they are mostly well away from sources of volatiles, such as later subduction zone activity. Both factors may be present to the north of the Sierra de la Ventana mountains, suggesting the possibility of kimberlites in southern Brazil and Uruguay, and the repetitive subduction along the Andean chain during the past billion years may also provide a volatile source, although Archaean upper mantle materials appear to be absent along this border.

Diamonds are a component of another intriguing feature of the South American Shield — the Roraima Formation — which forms the isolated mountain tops that were the locus for Conan Doyle's 'Lost World'. The basal conglomerates of this Precambrian sedimentary sequence are generally diamondiferous, and sedimentological studies (Gosh, Venezuela) show very clearly that the entire sequence came from the east. But what is the source of these sediments and where did the diamonds come from? The total original volume must have been at least 1 million km³. The age of the basement on which they lie is different in different areas

and it is generally thought that the base of the Roraima itself is strongly diachronous. There are major differences when the sediments are traced laterally, but the scale of this deposit makes such variations inevitable. The puzzle is why the similarities should persist over 1 million km². Sedimentological evidence clearly rules out the suggestion by Gaudett (New Hampshire) that local Venezuelan sources were of major importance. Gaudett and Olszewski demonstrated a consistent migration in the age of the formation of the basement rocks of southern Venezuela. Strontium isotopes showed that these Proterozoic granites and gneisses had little continental contamination, but their age decreased from ~ 1,850 Myr in the north-east to 1,550 Myr in the south-west. They suggested that this indicated a gradual migration of subduction activity up to a final suture zone close to the Venezuelan-Colombian border. What was important about Gaudett's study was that essentially similar techniques of whole-rock and zircon analyses were used. Many problems in understanding the geological evolution of South America arise from the wide variety of techniques that have been applied, each of which often represents the closure of a different geochemical system. Palaeomagnetic results, for example, from this and other Precambrian terrains, often assume that the radiometric age,

irrespective of the method, represents the age of the magnetization. This may be true for some K-Ar methods, but will generally not be the case for other techniques.

This radiometric dating problem is particularly important, on a world scale, in any attempt to understand the change from Archaean to Proterozoic tectonic conditions. There seems to be good evidence from virtually all other shield areas that Archaean tectonic conditions ceased between 2.5 and 2.7 billion years ago (Moorbath *Phil. Trans. R. Soc. A* **288**; 401, 1978). In northern Venezuela, the Imataca Province has an age of 3.1–3.3 billion years, but the oldest age, so far, from the Guyana Shield is 2.2–2.1 billion years (Gibbs, thesis, Harvard University, 1979) — indicating the persistence of 'Archaean' greenstone formation for some 500–600 Myr longer than in the rest of the world. Unless these dates have been significantly reset, they clearly require drastic modifications to be made to fit tectonic models that involve a world-wide change from Archaean- to Proterozoic-type tectonics over a period of some 200 Myr. Such critical problems can only be resolved by further work. It is unlikely that the next meeting (Brazil 1984) will provide all the answers, but the First Amazonian Symposium has certainly set the scene within which further work can be more effectively planned. □

REVIEW ARTICLE

Mechanisms of slow postsynaptic potentials

H. Criss Hartzell

Department of Anatomy, Emory University School of Medicine, Atlanta, Georgia 30322, USA

Classical views of synaptic transmission have to be modified to take account of slow postsynaptic potentials (p.s.ps), which are often mediated by novel ionic and molecular mechanisms and which are sometimes evoked by neuropeptides. Understanding slow p.s.ps will provide insights into the mechanisms of biological signal transduction and long-term signalling in the nervous system.

THE key to understanding synaptic transmission, neuronal integration and signalling is to discover how neurones recognize extracellular signals and transduce the signals into electrophysiological responses. Much of what we know about the recognition and transduction processes has been gathered from studies on the skeletal neuromuscular junction^{1,2,104} in which acetylcholine (ACh) released from the motor nerve terminals opens specific ionic channels in the membrane by binding to postsynaptic receptors. The response of the postsynaptic membrane to ACh is rapid^{3,4} (Fig. 1a); the conductance change is complete within 20 ms, the voltage response is slowed by the membrane time constant. Many synapses, however, do not work as rapidly. The homeostatic and autonomic functions of the nervous system, which operate with time courses of seconds or minutes, are not well served by fast synaptic responses lasting only milliseconds. For example, regulation of heart rate requires that transmitter action extend over several cycles of the beat⁵. Long transmitter action is often accomplished by slow synaptic responses which, in the heart, last several seconds (Fig. 1b) and in other targets may last tens of minutes (Table 1). Slow postsynaptic potentials (p.s.ps) are defined as responses which begin with a lag period of tens of milliseconds after transmitter application and are slower than the membrane time constant. Slow

p.s.ps, like fast p.s.ps, may be either excitatory (e.p.s.p.) or inhibitory (i.p.s.p.) depending on whether they increase or decrease the likelihood that the cell will reach threshold and initiate action potentials. Slow p.s.ps have attracted interest because they differ from fast p.s.ps in several fundamental ways. Slow p.s.ps: (1) are probably mediated by different molecular mechanisms from those mediating fast p.s.ps, (2) are often produced by novel ionic mechanisms which involve the closing of ionic channels or modulation of voltage-sensitive channels, and (3) are sometimes produced by non-classical transmitters, notably neuropeptides. The purpose of this review is to summarize these features of slow p.s.ps and to evaluate their possible mechanisms and roles. The interested reader is referred to several extensive reviews which have recently appeared^{6-12,105}.

Ionic mechanisms of slow p.s.ps

Many slow p.s.ps are mediated by conventional ionic mechanisms: a transmitter opens ionic channels and increases ionic conductance (g). For example, the slow response of Fig. 1b is due to an increase in the conductance of potassium (g_K)^{13,14}. Weight and Votava¹⁵ first proposed, however, that some slow

Table 1 Selected examples of slow postsynaptic potentials

Tissue	Species	Response	Δ Conductance	Latency	Duration	Transmitter	Refs
Sympathetic neurones	Mammal	Slow i.p.s.p.	$\downarrow Na$, electrogenic?	100 ms	1–5 s	ACh, DA?	6–9
		Slow e.p.s.p.	$\downarrow K$, $\uparrow Na$, $\uparrow Ca$ (?)	100–400 ms	10–60 s	ACh	6–9
	Frog	Slow i.p.s.p.	$\uparrow K$	100 ms	1–5 s	ACh	6–10, 103
		Slow e.p.s.p.	$\downarrow K$, $\uparrow Na$, $\uparrow Ca$ (?)	100–400 ms	10–60 s	ACh	6–10, 18
		Late-slow e.p.s.p.	$\downarrow K$, $\uparrow ?$	1–5 s	10–30 min	LHRH	29–31
Parasympathetic neurones	Mudpuppy	Slow i.p.s.p.	$\uparrow K$	30–100 ms	2–20 s	ACh	21
Mesenteric neurones	Guinea pig	Slow e.p.s.p.	$\uparrow$ (?)		2–60 s	Sub. P?	9, 34
Myenteric neurones	Guinea pig	Slow e.p.s.p.	$\downarrow K$	0.5–5 s	5–90 s	Sub. P, 5-HT	9, 33, 94
		Slow i.p.s.p.	$\uparrow K$	0.5–1 s	2–40 s	Non-cholinergic catecholamine	33
Submucous plexus neurones	Guinea pig	Slow i.p.s.p.	$\uparrow K$				95
L-Ped-3	<i>Planorbis</i>	Slow i.p.s.p.	$\uparrow K$	100 ms	5–10 ms	ACh	48
Cortical neurones	Cat	Slow e.p.s.p.	$\downarrow K$			ACh	96
L14 motor neurones	<i>Aplysia</i>	Slow e.p.s.p.	$\downarrow K$		10 s–3 min		93
Medial cells Neurones	<i>Aplysia</i>	Slow e.p.s.p.	$\downarrow K$	200–500 ms	5–20 s	ACh	38, 52
	<i>Helix</i>	Slow e.p.s.p.	$\downarrow K$		10 s	5-HT	39
		Slow i.p.s.p.	$\downarrow K$, $\downarrow Na$		20 s	5-HT	39
Cardiac muscle	Vertebrate	Slow i.p.s.p.	$\uparrow K$	100–300 ms	2–30 s	ACh	5, 21, 40–43
Smooth muscle	Vertebrate	Slow e.p.s.p.	$\uparrow K$, $\uparrow Na$, $\uparrow Ca$?	100–400 ms	2–30 s	ACh	11, 44, 45
Salivary gland	Cockroach	Slow i.p.s.p.	$\uparrow K$	1 s	10–20 s	DA	46
Submaxillary gland	Cat	Biphasic?	$\uparrow Na$, $\uparrow K$	500 ms	2–3 min	ACh	97
Parotid gland	Cat	Slow e.p.s.p.	$\uparrow Na$, $\uparrow K$	1–3 s	Minutes	Catecholamine (β)	97, 98
		Slow i.p.s.p.	?	300 ms	Seconds	ACh	19, 22, 97, 98
Pancreatic acinar cells	Cat	Hyperpolarization	$\uparrow Na$, $\uparrow K$	100–300 ms	2–3 min	ACh	20, 22, 97

Abbreviations: ACh, acetylcholine; DA, dopamine; LHRH, luteinizing hormone releasing hormone; 5-HT, serotonin.

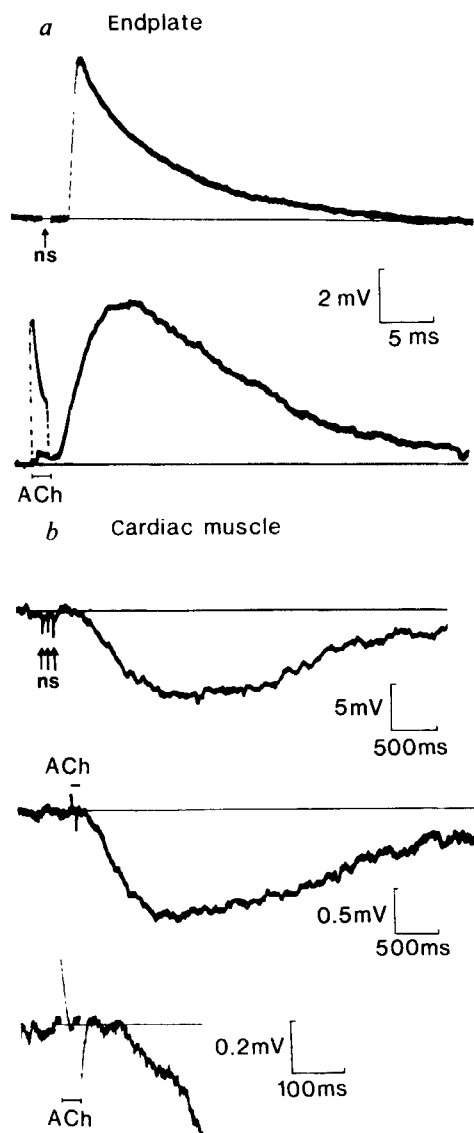


Fig. 1 Fast and slow p.s.ps. *a*, Response of frog skeletal muscle (cutaneous pectoris) to ACh. Top trace: the motor nerve was stimulated at the arrow (ns = nerve stimulation) with a 5-V, 0.5-ms duration pulse. The response of the muscle was recorded by an extracellular electrode positioned on the postsynaptic membrane. This type of recording measures the voltage drop produced by the synaptic current flowing across the resistance of the extracellular synaptic cleft. The time course of the response is virtually identical to that of the synaptic current. The time between the stimulus and the onset of the response is due to conduction time of the nerve and synaptic delay. Bottom trace: ACh was applied ionophoretically from a micropipette in a 1-ms pulse. The response is recorded with an intracellular voltage-recording electrode. The current pulse used to eject the ACh is superimposed on the intracellular record. The response begins during the 1-ms pulse of ACh. *b*, Response of muscle in sinus venosus of frog heart to ACh. Top trace: response to vagus nerve stimulation. Three shocks (arrows, ns) were delivered to the vagus nerve at 20 s⁻¹ and the response recorded with an intracellular electrode. Middle trace: response to a 50-ms duration pulse of ACh released from an iontophoretic pipette. Bottom trace: same as middle trace on a faster time base and at higher amplification to illustrate latency.

p.s.ps are produced by the closing of ionic channels which are normally open in the absence of transmitter. This suggestion was based on the finding that membrane resistance, monitored by the voltage deflection produced by a pulse of current injected into the cell, increased during the depolarizing slow e.p.s.p. in bullfrog sympathetic neurones^{15,16}. In addition, the reversal potential of the slow e.p.s.p. seemed to be in the direction of E_K because the amplitude of the slow e.p.s.p. usually decreased^{15,16} or reversed polarity^{15,17} with hyperpolarization. It was reasoned that the slow e.p.s.p. was produced by a decrease in g_K , as

follows. At membrane potentials positive to E_K , K^+ ions diffuse down their concentration gradient out of the cell. When transmitter turns off this resting outward current, the cell depolarizes towards the equilibrium potentials of the other permeant ions, mainly sodium.

Recent voltage-clamp studies have clearly demonstrated that the slow e.p.s.p.¹⁸ and the late-slow e.p.s.p.¹⁹ in bullfrog sympathetic ganglia involve turn-off of a unique K^+ current ('M-current'). An interesting feature of the M-current is that the channels are sensitive to membrane potential as well as to transmitter¹⁸. Near resting potential (-30 to -60 mV) the M-current channels are open, but they are closed at membrane potentials negative to -60 mV. In contrast, the conductance of channels responsible for fast p.s.ps, such as ACh-activated channels at the motor endplate, are not affected by membrane potential, although channel open-time is somewhat affected by voltage²⁰.

Voltage sensitivity is not restricted to p.s.ps produced by conductance decreases and may be a common property of slow p.s.ps. For example, a muscarinic slow i.p.s.p. produced by increased g_K in mudpuppy parasymphathetic neurones is voltage sensitive²¹. In the range of -100 mV to -60 mV, the ACh-induced increase in g_K is constant, but thereafter it decreases, until above -40 mV the channels are inactivated. This suggests that excitatory transmitters, including ACh acting on nicotinic receptors on the same neurone, have the power to inactivate and override the muscarinic slow i.p.s.p. With both the slow e.p.s.p. (above) and the slow i.p.s.p., the conductance change produced by transmitter decreases as the membrane potential moves away from the reversal potential of the response (that is, as the driving force of the response increases). Thus, transmitter will be most effective at a membrane potential between the reversal potential of the response and the membrane potential where channels are inactivated.

The ionic mechanisms of many slow p.s.ps have been difficult to elucidate. For example, because M-channels are inactivated at membrane potentials negative to -60 mV, accurate determination of the reversal potential of the slow e.p.s.p. has been difficult. Most investigators find that the amplitude of the slow e.p.s.p. decreases with hyperpolarization^{15,17,23,24}, as expected for a pure g_K decrease with reversal potential near E_K . However, a few investigators report that hyperpolarization increases the amplitude of the slow e.p.s.p. and that the slow e.p.s.p. is accompanied by increases in membrane conductance in some cells²³⁻²⁶. To explain these discrepancies, Kuba and Koketsu⁶ have proposed that the slow e.p.s.p. is produced both by a decrease in g_K and by increases in conductance of Na^+ and Ca^{2+} . When the cell is hyperpolarized below -60 mV, conductance increases are more noticeable because M-current is inactivated and driving forces on Na^+ and Ca^{2+} conductances increase. The ratio of the decrease and increases in conductance apparently varies in different cells⁶. Possibly, discrete populations of cells exist with different ionic mechanisms subject to different control, mediating their slow e.p.s.ps. Likewise, the ionic mechanisms of both the late-slow e.p.s.p. and the slow i.p.s.p. in sympathetic ganglia are confusing^{9,27,28,103}.

Peptides as neurotransmitters producing slow p.s.ps

The neuroactive peptides in peripheral nerves and ganglia may be responsible for slow p.s.ps. Good evidence for this hypothesis comes from Jan *et al.*^{27,29,30} working on the role of luteinizing hormone releasing hormone (LHRH) in the late-slow e.p.s.p.³¹ of bullfrog sympathetic neurones. Stimulation of the 8th and 9th spinal nerves in the presence of cholinergic antagonists produces a depolarization which may last 10–20 min, is caused partly by a decrease in g_K and is mimicked by the application of LHRH and certain analogues. Immunoreactive LHRH is concentrated in the preganglionic axons which synapse on the sympathetic neurones³⁰. In view of this evidence and the fact that LHRH is released from the ganglion in a calcium-dependent manner

when the ganglion is depolarized with high K^+ , LHRH meets the minimal criteria for a transmitter.

Peptides may also be transmitters in other ganglia. Substance P has been identified in nerve terminals of the guinea pig myenteric plexus³². Application of the peptide mimics the effects of stimulation of the nerves, producing a slow, decreased-conductance e.p.s.p. in the myenteric neurones³³. Substance P may also mediate a slow e.p.s.p. in mesenteric neurones in the guinea pig^{9,34}. Other peptides have been shown to produce membrane potential changes or modulate synaptic potentials in vertebrate autonomic ganglia and in invertebrate ganglia, but little direct evidence exists to support their roles as transmitters responsible for slow p.s.ps.

Molecular mechanisms of slow p.s.ps

The sequence of events leading to neurotransmitter-induced membrane conductance changes can be divided arbitrarily into four steps: (1) diffusion of transmitter to the receptive membrane, (2) binding of transmitter to receptor, (3) opening or closing of the ionic channel, which is presumed to be due to a conformational change in the channel, and (4) the relaxation of the channel to its initial closed or open conformation. In fast synaptic responses, such as the endplate current, the entire process is completed within tens of milliseconds³. Diffusion of ACh from the nerve terminal to the receptors takes place in less than 50 μ s, and the binding of ACh to receptors and channel opening are largely completed within 300 μ s (ref. 3). The time course of decay of the endplate current is due to the first-order closing of channels, which have exponentially distributed open times with a mean equal to the time constant of endplate current decay^{20,35-37}. Which of these steps is rate-limiting for slow p.s.ps? Because the time course of many nerve-evoked slow p.s.ps is mimicked by short iontophoretic pulses of transmitter^{3,11,21,38-46}, the slowness of these responses is almost certainly a postsynaptic phenomenon; but the possibility remains that very long-lasting p.s.ps produced by neuropeptides may be due in part to prolonged transmitter release¹⁰⁵.

Diffusion. The time course of slow p.s.ps is much slower than free diffusion in bulk solution^{5,21,40,42,43,46,47}, but the onset of slow p.s.ps might be slowed by diffusion barriers restricting access of transmitter to receptors^{11,12,40,41,47}. A variety of experiments argue against such a mechanism. (1) At several slow synapses no morphological diffusion barrier has been identified. For example, in cardiac muscle, where ACh produces a slow i.p.s.p. (Fig. 1b), receptors are diffusely distributed over the entire surface of the muscle and are not localized to clefts or other regions which appear inaccessible to transmitter⁴¹. (2) The same transmitter can produce both fast and slow p.s.ps in the same neurone. In parasympathetic neurones of the mudpuppy²¹, nicotinic ACh receptors mediate a fast e.p.s.p. which is complete in 100 ms, whereas muscarinic ACh receptors mediate a slow i.p.s.p. which is delayed 100 ms or more and lasts several seconds. Both nicotinic and muscarinic receptors are located on the neuronal soma. It seems unlikely that diffusion of ACh to muscarinic receptors is 100–1,000 times slower than its diffusion to nicotinic receptors. (3) The onset and latencies of slow p.s.ps are much more temperature sensitive than would be expected if diffusion were rate-limiting^{12,21,40,41,46,48,49}. Temperature coefficients (Q_{10}) as high as 10–12 have been reported. (4) Diffusion barriers should be overcome by large doses of transmitter. Latencies of slow p.s.ps, however, are rather insensitive to transmitter dose^{40,41}. (5) Nargeot *et al.*⁵⁰ have recently developed a technique to activate receptors in a way which virtually eliminates diffusion time of agonist to receptor. Tissue is equilibrated in agonist and a photosensitive antagonist. The antagonist can then be rapidly inactivated (<1 ms) by a flash of light. This exposes previously blocked receptors to agonist, which has been equilibrated with the tissue. In atrial muscle the increase in the ACh-induced K^+ current, caused by rapid inactivation of the antagonist *trans*-3-3'-bis-(α -(trimethylammonium)methyl)azobenzene in the presence of ACh,

has the same time course as the response produced by iontophoretic application of ACh.

Binding of transmitter to receptor. Hill-Smith and Purves⁴⁰ have calculated that the kinetics of transmitter binding can explain the decay of the slow p.s.p. In tissues where receptors are diffusely distributed⁴¹ the slow decay of p.s.ps might easily be explained by diffusion of transmitter along the receptive membrane and repeated binding to receptors. The quantitative effect of binding on the time course of the p.s.p., however, depends on the on- and off-rates of transmitter-receptor interaction. Kinetics of labelled agonist binding has not been published, but dissociation of muscarinic agonists is slow enough for equilibrium binding to be measured in an assay⁵¹ in which membranes are washed several times with buffer to remove free radiolabelled ligand, a process which requires several seconds. This suggests that kinetics of dissociation of agonist could explain the decay of slow p.s.ps.

This hypothesis is supported by the finding that the time course of the muscarinic hyperpolarization in cardiac muscle reflects the affinity of agonist binding to receptor. For example, Birdsall *et al.*⁵¹ have shown that oxotremorine binds to low- and high-affinity sites in brain with 10- and 21-fold higher affinity, respectively, than does carbamylcholine (CCh). In frog cardiac muscle, oxotremorine is 66 times more potent in competing for ³H-QNB (quinuclidinyl benzilate) binding than is CCh⁴¹. The hyperpolarization produced by iontophoretic application of oxotremorine is considerably slower (half decay time, $t_{1/2} = 5.1$ s) than the hyperpolarization produced by CCh ($t_{1/2} = 1.2$ s) (Fig. 2). The finding that oxotremorine has a higher affinity for receptor and produces a slower hyperpolarization favours the idea that the kinetics of agonist dissociation is partly responsible for the slowness of this p.s.p.

Although the decay of some slow p.s.ps may be accounted for by slow transmitter dissociation, the latency of slow p.s.ps is not easily explained⁴⁰ by the kinetics of agonist binding unless several steps occur in series between agonist binding and channel opening^{12,21,52}. These steps could involve the production of second messengers or cooperative interaction of several receptors in the opening of a channel.

Activation of conductance changes. The channels responsible for some slow p.s.ps may have very long open-times. Marty and Ascher⁵² have shown that an ACh-induced synaptic current in an *Aplysia* neurone relaxes very slowly in response to voltage jumps and have concluded that the mean channel open-time is several seconds.

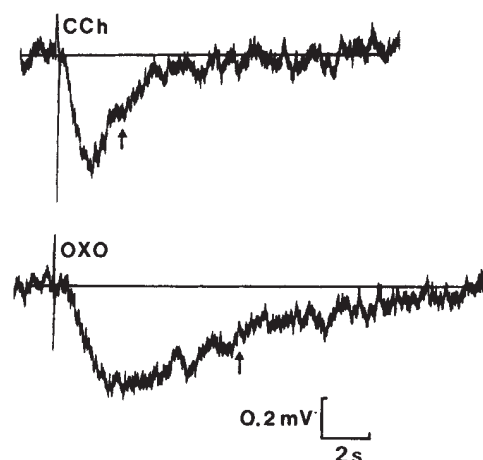


Fig. 2 Time course of responses of frog sinus venosus to muscarinic agonists. Upper trace: response to a 50-ms duration iontophoretic pulse of carbamylcholine (CCh). Bottom trace: response to a 5-ms duration iontophoretic pulse of oxotremorine (oxo). The pipette contained 0.5 M oxotremorine neutralized to pH 8 with HCl. Responses to both CCh and oxotremorine were on the linear portion of the dose-response curve. This was important because the time course of responses to muscarinic agonists applied iontophoretically is dose dependent⁴⁶. Arrows indicate time at which response declines to half its peak amplitude. Time from peak to half-decline time is $t_{1/2}$.

On the other hand, a mean channel open-time of ~ 100 ms for the ACh-activated K^+ channels in rabbit sinoatrial node has been inferred from analysis of ACh-induced noise^{53,54} and from relaxation of the ACh-induced current⁵⁵. Channel open-times vary somewhat with temperature, membrane potential and ACh concentration⁵⁴. Because mean channel open-time is short relative to the decay (several seconds) of the iontophoretic response, the first-order closing of channels with lifetimes of ~ 100 ms cannot be rate-limiting and some channels must be activated during the decay of the p.s.p. Repeated channel opening could be due to slow dissociation of agonist from the receptor or repeated agonist binding.

Channels which have been studied by single-channel recording usually have two states, open and closed, and transitions between these states are rapid³⁷. The onset of the slow p.s.p., therefore, probably reflects a slow increase in the fraction of open channels after transmitter binding rather than a slow transition of individual channels from one state to another. If transmitter binding instantaneously alters the opening (β) and closing (α) rate constants of the channel, the fraction of open channels will change exponentially with a time constant of $(\alpha' + \beta')^{-1}$, where α' and β' are the new rate constants. The estimates for α (10.1 s^{-1}) and β (12.0 s^{-1}) for the ACh-induced K^+ current in rabbit sinoatrial node⁵⁴ predict a time constant of ~ 45 ms, which is significantly less than what is observed for the onset of the response. Thus, factors other than channel kinetics probably contribute to the slowness of the onset (but compare ref. 56). The slow onset could be explained if channel opening is not mediated by direct transmitter-induced conformational changes of the channel but is caused indirectly by second messengers which are generated by agonist binding and build up slowly in the region of the channel.

It has been suggested that the second messengers mediating slow p.s.ps are cyclic nucleotides⁵⁷. In this scheme, transmitter activates an adenylate cyclase, which catalyses the production of cyclic AMP; cyclic AMP then activates a cyclic AMP-dependent protein kinase which phosphorylates the ionic channel. The channel is postulated to be open when phosphorylated and closed when dephosphorylated, or vice versa. An analogous mechanism may involve cyclic GMP⁵⁷. Good evidence exists for the involvement of cyclic AMP in mediating the increases in the action potential Ca^{2+} current produced by β -adrenergic agonists in cardiac muscle (reviewed in refs 12, 58) and by serotonin in an *Aplysia* neurone⁵⁹⁻⁶⁷. In both examples, (1) transmitter causes an increase in cyclic AMP production in the ganglia, (2) extracellular application of cyclic AMP derivatives or intracellular injection of cyclic AMP causes an increase in the Ca^{2+} current, and (3) agents, such as phosphodiesterase inhibitors, which increase cyclic AMP levels in the target, potentiate the electrophysiological effect of the transmitter. In the *Aplysia* neurone the electrophysiological action of serotonin correlates with increased ^{32}P incorporation into a 137,000 molecular weight membrane protein⁶¹. Further, intracellular injection of the purified catalytic subunit of cyclic AMP-dependent protein kinase mimics the electrophysiological effects of serotonin⁶⁰.

At other synapses the role of cyclic nucleotides is unresolved. For example, the concept that cyclic AMP mediates slow p.s.ps was advanced largely by the work of Greengard⁵⁷ and co-workers on the slow i.p.s.p. in neurones of sympathetic ganglia. Although impressive evidence has been amassed in support of this idea, this hypothesis has recently been challenged by several laboratories. The evidence has been evaluated by others^{9,10,28,62,63} and is summarized in Table 2. Perhaps the most serious problem is that only one laboratory^{64,65} of several which have tried^{10,63,66-74} has succeeded in mimicking the hyperpolarizing effect of the presumed transmitter, dopamine, by extracellular or intracellular application of cyclic AMP or cyclic AMP analogues. Furthermore, in the positive experiments^{64,65} no evidence was presented to show that the hyperpolarizations produced by cyclic AMP and dopamine occurred by identical ionic mechanisms. Another potential problem is that the transmitter which produces the slow i.p.s.p. may not be

dopamine, but ACh^{9,28,75,76,103}. If the natural transmitter is not dopamine, the basic assumption of many of the tests has been invalidated. Another controversy concerns whether the slow e.p.s.p. in sympathetic neurones is mediated by cyclic GMP⁷⁷⁻⁸⁰. Although evidence suggests that cyclic AMP may mediate membrane changes at some synapses, proof of this hypothesis must await further tests, including demonstration that transmitter-induced changes in membrane conductance are accompanied by cyclic AMP-dependent phosphorylation of appropriate membrane proteins.

Speculation about other second messengers

Second messengers are attractive mechanisms for explaining certain features of slow p.s.ps described in this review. As cyclic nucleotides may be involved in only certain p.s.ps, I will consider several other second messenger systems which could possibly have a role in mediating slow p.s.ps.

Information has accumulated that certain neurotransmitters, including those acting on α -adrenergic and muscarinic receptors, and some neuropeptides stimulate phosphatidylinositol breakdown to diacylglycerol and phosphatidic acid⁸¹; this has gained possible significance by the finding that phosphatidic acid can function as a calcium ionophore in artificial membranes⁸². Thus, phosphatidic acid produced by phosphatidylinositol breakdown in response to transmitter might act as a calcium pore in cell membranes. This suggestion is supported by the finding that exogenously applied phosphatidic acid produces contraction in smooth muscle cells⁸³. Further, the ACh-induced increase in the Ca^{2+} -activated K^+ permeability of parotid gland cells is mimicked by application of phosphatidic acid and calcium⁸⁴. In neutrophils, however, phosphatidylinositol breakdown seems to be a consequence and not a cause of Ca^{2+} influx⁸⁵.

The production of diacylglycerol as a result of phosphatidylinositol breakdown may also act as a second messenger in activating a novel protein kinase ('C-kinase') recently discovered by Nishizuka and co-workers⁸⁰. C-kinase could function to alter membrane permeability in an analogous way to that proposed for cyclic nucleotide-dependent protein kinases. It requires calcium and phosphatidylserine for its enzymatic activity⁸⁶. Enzyme activity is markedly stimulated by diacylglycerol^{187,88}, which increases the affinity of C-kinase for Ca^{2+} about 20-fold. Nishizuka has proposed that certain neurotransmitters activate membrane phospholipase C, which catalyses phosphatidylinositol breakdown to diacylglycerol. Diacylglycerol activates C-kinase in the presence of the low

Table 2 Mediation of the slow i.p.s.p. by cyclic AMP

Evidence in favour	Selected refs
Preganglionic nerve stimulation or dopamine application causes an increase in cyclic AMP levels in ganglion	64, 99-101
Increase in cyclic AMP is localized to ganglion cell soma	102
Dopamine-sensitive adenylate cyclase is present in ganglion	100
Methylxanthine phosphodiesterase inhibitors potentiate slow i.p.s.p.	63-66, 101
Extracellular application of cyclic AMP produces a hyperpolarization	64, 65
Evidence against	
Dopamine is less effective than preganglionic nerve stimulation or cholinergic agonists in producing increases in cyclic AMP	67, 101
Phosphodiesterase inhibitors which are not methylxanthines do not potentiate the slow i.p.s.p.	10, 63, 67
Methylxanthines have electrophysiological effects which differ from the effects of dopamine or nerve stimulation	10, 63, 66, 68
Extracellular application of cyclic AMP does not produce a hyperpolarization	10, 63, 66-72
Intracellular injection of cyclic AMP does not produce a hyperpolarization	10, 68, 71, 73, 74
Transmitter which mediates the slow i.p.s.p. may not be dopamine	9, 28, 75, 76, 103

intracellular concentrations of Ca^{2+} . No direct evidence links C-kinase with slow p.s.ps; however, the relationship of this enzyme to phospholipid turnover, which has been puzzling for many years, and the very high concentration of C-kinase in brain make the involvement of this enzyme an attractive possibility.

Phospholipids may also effect slow p.s.ps in another way. Axelrod and colleagues⁸⁹ have demonstrated that activation of β -adrenergic receptors in several systems can stimulate methyltransferases which catalyse methylation of phosphatidylethanolamine to phosphatidylcholine in the membrane. Phospholipid methylation may mediate ligand-induced changes in membrane permeability⁹⁰ (but see ref. 91). In mast cells, IgE binding to specific cell-surface receptors stimulates phospholipid methylation, calcium influx and histamine release⁹⁰. Inhibitors of the methyltransferases reduce phospholipid methylation by 90% and calcium influx by 85% (ref. 90). The mechanism by which phospholipid methylation affects calcium permeability is not known, although changes in membrane fluidity seem to be involved. Whether such a mechanism actually has a role in slow p.s.ps remains to be discovered.

Functional consequences of slow p.s.ps

The process of neuronal integration involves spatial and temporal summation of p.s.ps impinging on the neurone. Temporal summation of fast p.s.ps which last milliseconds requires a high degree of synchrony of converging inputs. Therefore perhaps the most important and obvious function of slow p.s.ps is to increase the length of time a synapse is active and to maximize the likelihood of temporal interaction of responses initiated at different times⁵. In addition, decreased conductance p.s.ps are particularly effective in modulating the efficacy of other inputs. For example, in sympathetic ganglia the ability of fast e.p.s.ps to initiate action potentials is increased during the late-slow e.p.s.p.^{27,92} due to (1) increased membrane resistance resulting in a greater voltage deflection for the same synaptic

current provided by the fast e.p.s.p., (2) summation of fast and late-slow e.p.s.ps to produce a greater depolarization, and (3) a lowered threshold for action potential generation. Activation of the late-slow e.p.s.p. can modulate the efficacy of fast e.p.s.ps from another input for up to 30 min. In addition, decreased conductance p.s.ps increase the membrane space constant and thus increase the area of the cell over which synaptic responses can interact.

The idea that such long-lasting heterosynaptic modulation has a role in behavioural responses has been demonstrated by Carew and Kandel⁹³. A decreased conductance e.p.s.p. in *Aplysia* increases the effectiveness of a previously ineffective stimulus in initiating inking behaviour. Noxious stimulation of the head of *Aplysia* produces both a fast e.p.s.p. and a slow, decreased-conductance e.p.s.p. in the motor neurones which evoke inking. Inking is usually triggered by the fast e.p.s.ps. The slow e.p.s.p. increases the likelihood that a stimulus to the siphon, which normally produces subthreshold e.p.s.ps in the motor neurone, will produce inking if siphon stimulation occurs during the slow e.p.s.p. In addition, the slow e.p.s.p. increases spatial summation of fast p.s.ps by increasing membrane resistance and space constant and the electrical coupling between motor neurones.

The abilities of neurotransmitters to regulate channels which are also gated by voltage and to modulate the effectiveness and summation of other synaptic events are important features associated with slow p.s.ps. Understanding these features will surely provide insight into mechanisms of neuronal integration and synaptic plasticity.

Thanks to Joanna Webb, Vicki Shadix, Jan Zwemer-Collins, Doug Carr, Russell Fricke, Stephen Schneider, John Wood and John Scott for help with and comments on the manuscript, and to Fang for encouragement. Research in my laboratory is funded by NIH Research Career Development Award HL00435 and research grants HL21195 and HL25090.

- Colquhoun, D. A. *Rev. Pharmac.* **15**, 307-325 (1975).
- Steinbach, J. H. & Stevens, C. F. in *Handbook of Frog Neurobiology* (eds Llinas, R. Precht, W.) 33-92 (Springer, Berlin, 1976).
- Katz, B. *The Release of Neural Transmitter Substances* (Thomas, Springfield, 1969).
- Kuffler, S. W. & Yoshikami, D. *J. Physiol., Lond.* **244**, 703-730 (1975).
- Purves, R. D. *Nature* **261**, 149-151 (1976).
- Kuba, K. & Koketsu, K. *Prog. Neurobiol.* **11**, 77-169 (1978).
- Nishi, S. in *The Peripheral Nervous System* (ed. Hubbard, J. I.) 225-255 (Plenum, New York, 1974).
- Libet, B. *Fedn Proc.* **29**, 1945-1956 (1970).
- Dun, N. J. *Fedn Proc.* **39**, 2982-2989 (1980).
- Weight, F. F., Schulman, J. A., Smith, D. A. & Busis, N. A. *Fedn Proc.* **38**, 2084-2094 (1979).
- Bolton, T. D. *Physiol. Rev.* **59**, 606-718 (1979).
- Kehoe, J. S. & Marty, A. A. *Rev. Biophys. Bioengng* **9**, 437-465 (1980).
- Hutter, O. F. in *Nervous Inhibition* (ed. Florey, E.) 144-123 (Pergamon, New York, 1961).
- Garnier, D., Nargeot, J., Ojeda, C. & Rougier, O. *J. Physiol., Lond.* **274**, 381-396 (1978).
- Weight, F. F. & Votava, J. *Science* **170**, 755-758 (1970).
- Kobayashi, H. & Libet, B. *J. Physiol., Lond.* **208**, 353-372 (1970).
- Weight, F. F. & Votava, J. *Science* **172**, 503-504 (1971).
- Brown, D. A. & Adams, P. R. *Nature* **283**, 673-675 (1980).
- Adams, P. R. & Brown, D. A. *Br. J. Pharmac.* **68**, 353-355 (1980).
- Anderson, C. R. & Stevens, C. F. *J. Physiol., Lond.* **238**, 655-691 (1973).
- Hartzell, H. C., Kuffler, S. W., Stickgold, R. & Yoshikami, D. *J. Physiol., Lond.* **271**, 817-846 (1977).
- Kobayashi, H. & Libet, B. *Proc. natn. Acad. Sci. U.S.A.* **60**, 1304-1311 (1968).
- Kuba, K. & Koketsu, K. *Jap. J. Physiol.* **26**, 651-669 (1976).
- Nishi, S., Soeda, H. & Koketsu, K. *Life Sci.* **8**, 33-42 (1969).
- Libet, B. *Science* **172**, 503-504 (1971).
- Kobayashi, H. & Libet, B. *Life Sci.* **14**, 1871-1883 (1974).
- Jan, Y. N., Jan, L. Y. & Kuffler, S. W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5008-5012 (1980).
- Gallagher, J. P., Shinnick-Gallagher, P., Cole, A. E., Griffith III, W. H. & Williams, B. J. *Fedn Proc.* **39**, 3009-3015 (1980).
- Jan, Y. N., Jan, L. Y. & Kuffler, S. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1501-1505 (1979).
- Jan, L. Y., Jan, Y. N. & Brownfield, M. S. *Nature* **288**, 380-382 (1980).
- Nishi, S. & Koketsu, K. *J. Neurophysiol.* **31**, 109-121 (1968).
- Pearse, A. G. & Polark, J. M. *Histochemistry* **41**, 373-375 (1975).
- Johnson, S. M., Katayama, Y. & North, R. A. *J. Physiol., Lond.* **301**, 505-516 (1980).
- Dun, N. J. & Karczmar, A. G. *Neuropharmacology* **18**, 215-218 (1979).
- Katz, B. & Miledi, R. *J. Physiol., Lond.* **224**, 665-699 (1972).
- Neher, E. & Sakmann, B. *Nature* **260**, 799-802 (1976).
- Neher, E. in *Techniques in Cellular Physiology* (ed. Baker, P. F.) (Elsevier, Amsterdam, 1980).
- Kehoe, J. J. *Physiol., Lond.* **225**, 147-172 (1972).
- Gerschenfeld, H. M. & Paupardin-Tritsch, D. *J. Physiol., Lond.* **243**, 457-481 (1974).
- Hill-Smith, I. & Purves, R. D. *J. Physiol., Lond.* **279**, 31-54 (1978).
- Hartzell, H. C. *J. Cell Biol.* **86**, 6-20 (1980).
- Glitsch, H. G. & Pott, L. *J. Physiol., Lond.* **279**, 655-668 (1978).
- Pott, L. *Pflügers Arch. ges. Physiol.* **380**, 71-77 (1979).
- Bolton, T. B. *Br. med. Bull.* **35**, 275-283 (1979).
- Bennett, M. R. *Autonomic Neuromuscular Transmission* (Cambridge University Press, 1972).
- Blackman, J. G., Ginsborg, B. L. & House, C. R. *J. Physiol., Lond.* **287**, 67-80 (1979).
- Purves, R. D. *J. theor. Biol.* **65**, 327-344 (1977).
- Katchman, A. N., Ger, B. A. & Zeimal, E. V. *Gen. Pharmac.* **11**, 55-64 (1980).
- Creed, K. E. & McDonald, I. R. *J. Physiol., Lond.* **247**, 521-535 (1975).
- Nargeot, J. *et al. Biophys. J.* **33**, 15a (1981).
- Birdsall, N. J. M., Burgen, A. S. V. & Hulme, E. C. *Molec. Pharmac.* **14**, 723-736 (1978).
- Marty, A. & Ascher, P. *Nature* **274**, 494-497 (1978).
- Noma, A., Peper, K. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **381**, 225-262 (1979).
- Osterrieder, W., Noma, A. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **386**, 101-109 (1980).
- Noma, A. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **377**, 193-200 (1978).
- Osterrieder, W., Yang, Q.-F. & Trautwein, W. *Pflügers Arch. ges. Physiol.* (in the press).
- Greengard, P. *Science* **199**, 146-152 (1978).
- Tsien, R. W. *Adv. Cyclic Nucleotide Res.* **8**, 363-420 (1977).
- Klein, M., Shapiro, E. & Kandel, E. R. *J. exp. Biol.* **89**, 117-157 (1980).
- Castellucci, V. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 7492-7496 (1980).
- Paris, C. G., Castellucci, V. F., Kandel, E. R. & Schwartz, J. H. in *Cold Spring Harb. Conf. Cell Proliferation* **8**, in press (1981).
- McAfee, D. A. *et al. Fedn Proc.* **39**, 2997-3002 (1980).
- Libet, B. *Life Sci.* **24**, 1043-1058 (1979).
- McAfee, D. A. in *SIF Cells* (ed. Eranko, O.) 132-142 (U.S. Govt Printing Office, 1975).
- McAfee, D. A. & Greengard, P. *Science* **178**, 310-312 (1972).
- Dun, N. J. & Karczmar, A. G. *J. Pharmac. exp. Ther.* **202**, 89-96 (1977).
- Brown, D. A., Caulfield, M. P. & Kirby, P. J. *J. Physiol., Lond.* **290**, 441-451 (1979).
- Busis, N. A., Weight, F. F. & Smith, P. A. *Science* **200**, 1079-1081 (1978).
- Dun, N. J., Kaibara, K. & Karczmar, A. G. *Science* **197**, 778-779 (1977).
- Akasu, T. & Koketsu, K. *Br. J. Pharmac.* **60**, 331-336 (1977).
- Kobayashi, H., Hashiguchi, T. & Ushiyama, N. S. *Nature* **271**, 268-270 (1978).
- Busis, N. A., Schulman, J. A., Smith, P. A. & Weight, F. F. *Br. J. Pharmac.* **62**, 378P-379P (1978).
- Gallagher, J. P. & Shinnick-Gallagher, P. *Science* **198**, 851-852 (1977).
- Weight, F. F., Smith, P. A. & Schulman, J. A. *Brain Res.* **158**, 197-202 (1978).
- Weight, F. F. & Padjen, A. *Brain Res.* **55**, 219-224 (1973).
- Weight, F. F. & Padjen, A. *Brain Res.* **55**, 225-228 (1973).
- Weight, F. F., Petzold, G. & Greengard, P. *Science* **186**, 942-944 (1974).
- Hishiguchi, T., Ushiyama, N. S. & Kobayashi, H. *Nature* **271**, 267-268 (1968).
- Dun, N. J., Kaibara, K. & Karczmar, A. G. *Neuropharmacology* **16**, 715-717 (1977).
- Dun, N. J., Kaibara, K. & Karczmar, A. G. *Brain Res.* **150**, 658-661 (1978).
- Michell, R. *Biochim. biophys. Acta* **415**, 81-147 (1975).
- Tyson, C. A., Vande Zande, H. & Green, D. E. *J. Biol. Chem.* **251**, 1326-1332 (1976).
- Salmon, D. M. & Honeyman, T. W. *Nature* **284**, 344-345 (1980).
- Putney, J. W., Weiss, S. J., Van De Walle, C. M. & Haddas, R. A. *Nature* **284**, 345-347 (1980).
- Cockcroft, S., Bennett, J. P. & Gomperts, B. D. *Nature* **288**, 275-277 (1980).
- Takai, Y., Kishimoto, A., Iwasa, Y., Kawahara, Y., Mori, T. & Nishizuka, Y. *J. Biol. Chem.* **254**, 3692-3695 (1979).
- Takai, Y., Kishimoto, A., Kikkawa, U., Mori, T. & Nishizuka, Y. *Biochem. biophys. Res. Commun.* **91**, 1218-1225 (1979).

88. Kishimoto, A., Takai, Y., Mori, T., Kikkawa, U. & Nishizuka, Y. *J. biol. Chem.* **255**, 2273–2276 (1980).
 89. Hirata, F. & Axelrod, J. *Science* **209**, 1082–1090 (1980).
 90. Ishizaka, T., Hirata, F., Ishizaka, K. & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1903–1906 (1980).
 91. Vance, D. E. & de Kruijff, B. *Nature* **288**, 277–278 (1980).
 92. Schulman, J. A. & Weight, F. F. *Science* **194**, 1437–1439 (1976).
 93. Carew, T. J. & Kandel, E. R. *Science* **192**, 150–153 (1976).
 94. Wood, J. D. & Mayer, C. J. *J. Neurophysiol.* **42**, 582–593 (1979).
 95. Hirst, G. D. S. & Silinsky, E. M. *J. Physiol., Lond.* **251**, 817–832 (1975).
 96. Krnjec, K., Pumain, R. & Renaud, L. *J. Physiol., Lond.* **215**, 247–268 (1971).

97. Petersen, O. H. *Physiol. Rev.* **56**, 535–577 (1976).
 98. Emmelin, N., Grampp, W. & Thesleff, P. *J. Physiol., Lond.* **302**, 183–195 (1980).
 99. McAfee, D. A., Schorderet, M. & Greengard, P. *Science* **171**, 1156–1158 (1971).
 100. Kebabian, J. W. & Greengard, P. *Science* **174**, 1346–1349 (1971).
 101. Kalix, P., McAfee, D. A., Schorderet, M. & Greengard, P. *J. Pharmac. exp. Ther.* **188**, 676–686 (1974).
 102. Kebabian, J. W., Bloom, F. E., Steiner, A. L. & Greengard, P. *Science* **190**, 157–159 (1975).
 103. Dodd, J. & Horn, S. *Nature* (submitted).
 104. Adams, P. R. *J. Membrane Biol.* **58**, 161–174 (1981).
 105. Kuffler, S. W. *J. exp. Biol.* **81**, 257–286 (1980).

ARTICLES

Graphite–magnetite aggregates in ordinary chondritic meteorites

Edward R. D. Scott, G. Jeffrey Taylor, Alan E. Rubin, Akihiko Okada* & Klaus Keil

Institute of Meteoritics and Department of Geology, University of New Mexico, Albuquerque, New Mexico 87131, USA

Aggregates consisting largely of submicrometre graphite and magnetite crystals have been identified in 14 unequilibrated and regolith breccia, ordinary chondrites. Like other primitive components, they formed before the meteorites accreted. Chondrite clasts rich in graphite–magnetite occur in three regolith breccias and are a new kind of unequilibrated ordinary chondrite. They were probably the source of graphite–magnetite aggregates in regolith breccias.

ORDINARY chondrites are believed to have formed from three major components which can easily be identified in the most unequilibrated chondrites^{1,2}: chondrules, metallic Fe,Ni, and fine-grained, Si-rich, opaque (and recrystallized) matrix—'Huss matrix'³. These components formed in different conditions before the accretion of the meteorites. We have discovered a new component in all three ordinary chondrite groups consisting largely of an aggregate of submicrometre sized grains of graphite and magnetite (Fe₃O₄). The composition and distribution of the graphite–magnetite component were studied by optical microscopy, electron-microprobe analysis and X-ray powder diffraction.

We have found the graphite–magnetite component in three types of occurrences: as abundant isolated inclusions in eight ordinary-chondritic, regolith breccias ('gas-rich' breccias), as the sole matrix in a new kind of unequilibrated chondrite that forms clasts in these regolith breccias, and together with Huss matrix in six unequilibrated ordinary chondrites. We suggest that this component formed by low-temperature, gas–solid reactions before the accretion of the meteorites. We also suggest that the isolated inclusions of graphite–magnetite in regolith breccias were derived from bodies composed of the new kind of chondrite that has graphite–magnetite as its sole matrix.

Experimental findings

Twenty isolated graphite–magnetite inclusions were found in the Dimmitt H chondrite; they range in diameter from 30 µm to 1 mm (Fig. 1a). Most graphite crystals are submicrometre in size, but some are as large as 5 µm and show the characteristic optical anisotropy, birefractance and softness of graphite. X-ray diffraction of a sample scratched from a millimetre-sized inclusion (Table 1) shows lines from graphite and a spinel, which was identified as magnetite by electron-probe analysis (Table 2). Table 1 also shows lines of aragonite (CaCO₃), but because CaCO₃ occurs in corrosion veins in the inclusions and we found low-Ca concentrations between these veins (Table 2), we believe the aragonite is a terrestrial contaminant (Dimmitt is a meteorite find).

Accurate electron-probe analyses of these inclusions were not possible because the crystals are smaller than the electron beam width (1–2 µm), and because of great uncertainties in the correction procedures for carbon analyses⁴. Figure 2 shows approximate Fe concentration plotted against the intensity of carbon K radiation in four Dimmitt inclusions. The data define a mixing curve between C and Fe₃O₄. Strong absorption of carbon X rays by Fe causes curvature of the mixing line. Table 2 shows mean compositions of graphite–magnetite inclusions in Dimmitt

Table 1 X-ray powder diffraction patterns of a graphite–magnetite inclusion in the Dimmitt chondrite*, and of graphite, magnetite and aragonite standards

Dimmitt inclusion†		Graphite‡ C (ref. 38)		Magnetite Fe ₃ O ₄ (ref. 39)		Aragonite CaCO ₃ (ref. 39)	
d(Å)	I/I ₀	d(Å)	I/I ₀	d(Å)	I/I ₀	d(Å)	I/I ₀
4.21‡	vw						
3.39	vs	3.37	10			3.396	10
3.28	w					3.273	5
2.956	vw			2.967	3		
2.700	w					2.700	5
2.520	m			2.532	10		
2.481	vw					2.481	3
2.383	vw					2.372	4
2.337	vw					2.341	3
2.106	w	2.136	4	2.099	2	2.106	2
2.026	vw	2.044	5				
1.977	w					1.977	7
1.882	vw					1.882	3
						1.877	3
						1.814	2
1.815	vw					1.742	3
1.740	vw						
1.689	vw	1.691	5				
1.605	vw			1.616	3		
1.477	vw			1.485	4		

* Using FeKα radiation.

† After treatment with 6M HCl the sample gave only graphite powder lines 3.39 (vs), 2.203 (w), 1.690 (vw), 1.288 (vw) and 1.155 (vw). vs, Very strong; w, weak; vw, very weak; m, medium.

‡ Probably the strongest line of goethite (α-FeOOH).

§ Spacings and intensities of graphite diffraction lines vary considerably.

* Permanent address: Institute of Physical and Chemical Research, Wako, Saitama, Japan.

and two other meteorites obtained with a 10- μ m wide electron beam. Even with such a broad beam, elemental concentrations vary considerably from place to place within a single inclusion: Fe usually varies by 10–50% relative from the mean, minor elements such as Si, Ni, S, Cr and Mg vary by ~50–100% relative, although their total is usually below 5 wt%. Carbon concentrations, which vary overall between 40 and 98 wt%, were estimated by difference assuming that S is in the form of FeS, the remaining Fe is in Fe₃O₄ and other elements are present as oxides. Analyses of graphite-magnetite in seven other chondrites all plot on or near the mixing curve shown in Fig. 2, and show comparable minor element concentrations.

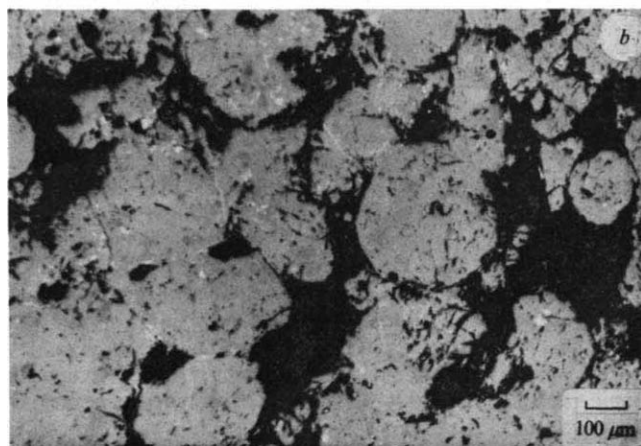
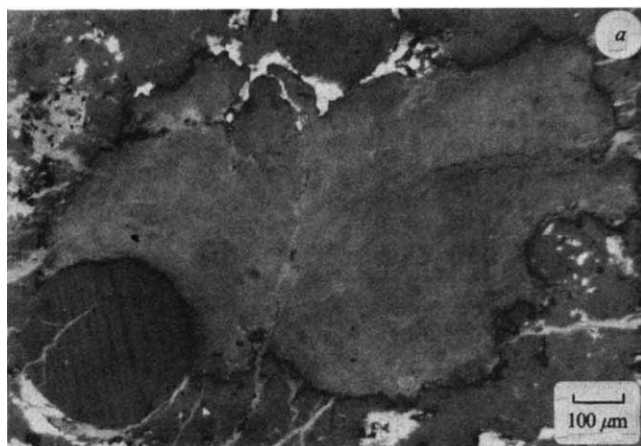


Fig. 1 Reflected-light photomicrographs of graphite-magnetite aggregates in ordinary chondrites. *a*, An irregular isolated graphite-magnetite inclusion in the Dimmitt regolith breccia. (Veins were produced by terrestrial weathering.) *b*, Part of the clast PV1 in the Plainview regolith breccia. This clast is a new kind of chondrite that contains chondrules in a graphite-magnetite matrix.

Isolated graphite-magnetite inclusions like those in Dimmitt were also found in six other H chondrites—Abbott, Elm Creek, Leighton, Plainview, Tysnes Island and Weston—and in one L chondrite—Grassland. All these meteorites contain abundant solar-wind gases^{5,6}, indicating that they are regolith breccias^{7,8} which formed by collisional mixing of meteorite material. We estimate that these chondrites contain between 10⁻² and 10⁻⁶ g per g of graphite-magnetite inclusions, and predict that other gas-rich ordinary chondrites, and perhaps gas-rich achondrites as well, contain similar concentrations of inclusions.

A new kind of unequilibrated ordinary chondrite containing even higher concentrations of graphite-magnetite was discovered in the form of four clasts in Plainview (PV1, which was

Table 2 Mean compositions (wt %)* of graphite-magnetite in Dimmitt, Plainview and Allan Hills A77011 chondrites determined by broad-beam electron-microprobe analysis (5–10 μ m beam diameter)

Meteorite	Dimmitt			Plainview	A77011‡
No. of analyses	Incl. 1†	Incl. 5	Clast DT1	Clast PV1	
Na	0.09	0.13	0.18	0.08	<0.05
Mg	<0.03	0.11	0.4	0.6	0.4
Al	<0.03	<0.03	0.15	0.08	0.06
Si	0.35	0.7	1.4	1.4	0.6
P	<0.03	<0.03	0.1	0.04	<0.03
S	0.11	0.4	0.5	0.25	0.25
Ca	0.09	0.06	0.15	0.4	0.09
Cr	<0.05	<0.05	0.2	0.5	<0.05
Fe	27	9	31	11	9
Ni	1.0	0.4	4	3	1.4
C by diff.	(60)	(84)	(47)	(75)	(83)

* Other elements measured: Cl ≤ 0.1%, K < 0.1%, Ti < 0.05%, Mn ≤ 0.1%, Co < 0.1% (0.6% in DT1), Cu < 0.2%, and Zn < 0.3%.

† Inclusion is shown in Fig. 1*a*, and plotted in Fig. 2 as inclusion 1.

‡ Specimen Allan Hills A77043, which like A77214 and several other Allan Hills specimens is paired with Allan Hills A77011.

described by Fodor and Keil⁹), Dimmitt (DT1 and DT2) and Weston (WN1). These four clasts, which are only 1–5 mm in diameter, contain 15–35 vol% of graphite-magnetite matrix, which fills much of the space between the chondrules (Fig. 1*b*). Like normal unequilibrated ordinary chondrites (UOCs), the clasts contain heterogeneous olivines and pyroxenes, olivines with high Ca concentrations, and well-defined chondrules containing glass¹. But they differ in having a graphite-magnetite matrix instead of the normal Huss matrix, and consequently have much higher C concentrations (7–17 wt%) than normal UOCs (0.2–1.0 wt%) and other meteorites^{10,11}. Metallic Fe, Ni abundances in the clasts seem to be lower than those of most UOCs. However, shock veining and terrestrial corrosion prevent accurate estimates of metal abundances.

We have discovered two UOCs that contain abundant graphite-magnetite in addition to the normal Huss matrix: Sharps (H3) and Allan Hills A77011 (L3). Sharps contains more graphite-magnetite than Huss matrix, whereas in A77011, the

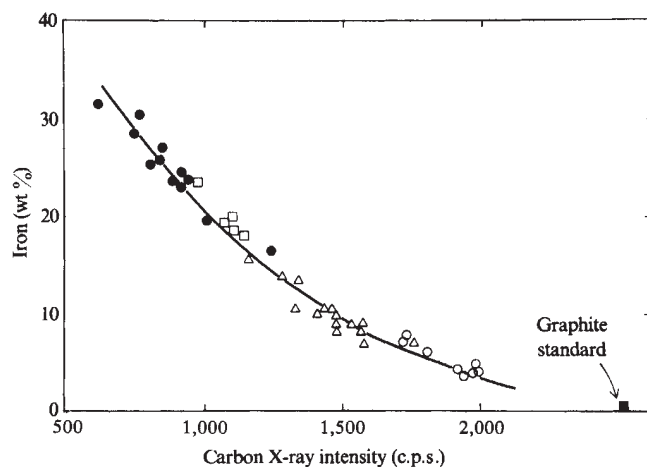


Fig. 2 Approximate Fe concentration in four Dimmitt inclusions (●, 1; □, 2; △, 3; ○, 4) plotted against the intensity of carbon X rays (K line, background corrected) measured simultaneously with the electron microprobe. The data define a mixing curve between pure carbon and an Fe-rich mineral, consistent with X-ray diffraction data, which indicate that the inclusions are largely composed of graphite and magnetite.

reverse is true. These two meteorites contain very heterogeneous silicates and have high bulk carbon concentrations^{12,13}, characteristics of the most unequilibrated ordinary chondrites^{1,3,14}. In both meteorites, tiny metallic Fe,Ni grains commonly occur within graphite-magnetite inclusions. Their morphology and distribution suggest that they formed by reduction of magnetite, possibly before the meteorites accreted. (Fredriksson *et al.*¹² also observed graphite in Sharps, but not magnetite, and suggested that metal formed by the reducing action of carbon.) Four other UOCs have been found which contain very minor graphite-magnetite (<1 vol%): Allan Hills A77278 (LL3), Allan Hills A77299 (H3), Dhajala (H3) and Bremervörde (H3). In these meteorites also, the graphite-magnetite is associated with Fe,Ni metal and distributed around chondrules. A full description of these clasts and chondrites will be given elsewhere.

Discussion

Although composed of very abundant elements, graphite and magnetite have been considered relatively rare minerals in ordinary chondrites¹⁵. Their aggregates may have escaped previous detection because in transmitted light they appear opaque like metal or unrecrystallized Huss matrix, whereas in reflected light they look grey like silicates. Here we discuss the origin and implications of graphite-magnetite and its occurrence in unequilibrated clasts and meteorites and in chondritic regolith breccias.

The presence of neon-E, which is almost pure ²²Ne (ref. 16), in a graphite-rich acid-resistant residue from Dimmitt¹⁷ suggests that the graphite-magnetite we observed might be the carrier of Ne-E and, hence, presolar in origin^{18,19}. (See also ref. 20 for views on the origin of Ne-E.) However, noble-gas analyses of samples from a large graphite-magnetite inclusion in Dimmitt and the clast PV1 failed to reveal Ne-E (ref. 21). Thus, we conclude that either Ne-E is heterogeneously distributed in these inclusions, or it is present in another phase in the residue. The abundance of submicrometre graphite, and perhaps magnetite also, in interstellar grains^{22,23} and the occurrence of Ne-E in carbon-rich phases in CI and CM carbonaceous chondrites²⁴⁻²⁶ provide weak evidence favouring a presolar source for the inclusions. If they are presolar, their abundance in the unequilibrated clasts implies that at the formation location of these clasts in the solar nebula, a significant fraction of the accreting material was interstellar dust and not solar nebula condensates. This would support Clayton's view²⁷ that temperatures were not high enough to vaporize interstellar grains at the formation location of meteorites.

Like Huss matrix, the graphite-magnetite component seems to have formed by gas-solid reactions before accretion of the meteorites. According to equilibrium calculations, in a gas of solar composition, graphite only forms at low temperatures and

pressures, less than 400 K and 10⁻⁸ atm (refs 28, 29). However, it is unlikely that equilibrium could be obtained in practice in these conditions. Alternatively, if the carbon abundance in the gas exceeds that of oxygen, graphite is stable over a wide range of temperatures and pressures and is the most abundant solid to form at equilibrium. Iron metal is stable at 10⁻⁴ atm below 1,400 K, but reacts to form FeS below ~600 K and Fe₃O₄ below ~400 K in systems with solar abundances^{22,30}. Thus graphite and iron are major condensates in C-rich environments, but if the graphite-magnetite inclusions equilibrated with a gas it is difficult to understand their low abundances of FeS and Si-rich minerals. Huss matrix, which has been considered to be a low-temperature condensate^{3,30}, also contains C and Fe₃O₄ (refs 31, 32) but in concentrations far lower than those of the inclusions we have described.

The graphite-magnetite component resembles the Huss matrix in its submicrometre crystal size, and circumstantial evidence suggests that both may be enriched in volatile elements^{3,21}. However, we have not observed any mixing of these components and conclude that they formed in very different conditions. We suggest that UOCs formed from four major components: chondrules, metallic Fe,Ni, Huss matrix, and graphite-magnetite. For unknown reasons, some UOCs failed to accrete Huss matrix and are entirely composed of chondrules and graphite-magnetite. We conclude that this material, represented by the four clasts, constitutes a new kind of unequilibrated chondrite.

Finally, we suggest that unequilibrated chondrites rich in graphite-magnetite were the source of the abundant graphite-magnetite inclusions in ordinary-chondritic regolith breccias. These chondrites may have been derived from subkilometre bodies that were too small to be metamorphosed by ²⁶Al heating. (Chondritic regolith breccias and unequilibrated chondrites experienced only mild metamorphism after compaction. Stronger heating might have reduced magnetite to metal.) These unequilibrated chondrites may also have been a source for the carbon (and volatile?) enrichments in the dark hosts of ordinary-chondritic regolith breccias^{33,34}. Neither carbonaceous chondrites nor normal UOCs seem capable of providing the observed volatile enrichments^{34,35}. A flux of bodies rich in graphite-magnetite in the Solar System might also account for some of the discrepancies between the reflectivities of asteroids and common chondrites³⁶. Many asteroids have flat reflectance spectra and low albedoes, characteristics that can be reproduced by the presence of a small admixture of opaques like graphite and magnetite³⁷.

We thank R. S. Clarke Jr, M. E. Lipschutz and the Antarctic Meteorite Working Group for the loan of meteorite samples, G. H. Conrad and J. S. Allen for technical assistance, R. Hutchison for comments, and B. Hudson and C. M. Hohenberg for noble-gas analyses. This work was supported by NASA grant NGL 32-004-064.

Received 27 January; accepted 1 April 1981.

- Dodd, R. T., Van Schmus, W. R. & Kaufman, D. M. *Geochim. cosmochim. Acta* **31**, 921-951 (1967).
- Wood, J. A. *Icarus* **2**, 152-180 (1963).
- Huss, G. R., Keil, K. & Taylor, G. J. *Geochim. cosmochim. Acta* **45**, 33-51 (1981).
- Reed, S. J. B. *Electron Microprobe Analysis*, 318-325 (Cambridge University Press, 1975).
- Schultz, L. & Kruse, H. *Nucl. Track Detect.* **2**, 65-103 (1975).
- Wasson, J. T. *Meteorites—Classification and Properties*, 97-104 (Springer, Berlin, 1974).
- Pellas, P. in *From Plasma to Planet* (ed. Elvius, A.) 65-92 (Wiley, New York, 1972).
- Rajan, R. S. *Geochim. cosmochim. Acta* **38**, 777-788 (1974).
- Fodor, R. V. & Keil, K. *Geochim. cosmochim. Acta* **40**, 177-189 (1976).
- Moore, C. B. & Lewis, C. F. *J. geophys. Res.* **72**, 6289-6292 (1967).
- Otting, W. & Zähringer, J. *Geochim. cosmochim. Acta* **31**, 1949-1960 (1967).
- Fredriksson, K., Jarosewich, E. & Nelen, J. in *Meteorite Research* (ed. Millman, P. M.) 155-165 (Springer, Berlin, 1969).
- Marvin, U. B. & Mason, B. *Smiths. Contr. Earth Sci.* **23**, 1-50 (1980).
- Sears, D. W., Grossman, J. N., Melcher, C. L., Ross, L. M. & Mills, A. A. *Nature* **287**, 791-795 (1980).
- Ramdohr, P. *The Opaque Minerals in Stony Meteorites* (Elsevier, Amsterdam, 1973).
- Jungck, M. H. A. & Eberhardt, P. *Meteoritics* **14**, 439-441 (1979).
- Niederer, F. & Eberhardt, P. *Meteoritics* **12**, 327-331 (1977).
- Black, D. C. *Geochim. cosmochim. Acta* **36**, 377-394 (1972).
- Clayton, D. D. *Nature* **257**, 36-37 (1979); *Space Sci. Rev.* **24**, 147-226 (1979).

- Sabu, D. D. & Manuel, O. K. *Proc. 11th Lunar Planet. Sci. Conf.* 879-899 (1980).
- Scott, E. R. D. *et al. Lunar planet Sci.* **12**, 955-957 (1981).
- Salpeter, E. E. *Astr. Astrophys.* **15**, 267-293 (1977).
- Knacke, R. F. in *Protostars and Planets* (ed. Gehrels, T.) 112-133 (University of Arizona, Tucson, 1978).
- Lewis, R. S., Alarits, L., Matsuda, J. & Anders, E. *Astrophys. J. Lett.* **234**, L165-L168 (1979).
- Eberhardt, P., Jungck, M. H. A., Meier, F. O. & Niederer, F. *Astrophys. J. Lett.* **234**, L169-L171 (1979).
- Whittaker, A. G., Watts, E. J., Lewis, R. S. & Anders, E. *Science* **209**, 1512-1514 (1980).
- Clayton, D. D. *Moon Planets* **19**, 109-137 (1978).
- Sears, D. W. *Earth planet. Sci. Lett.* **41**, 128-138 (1978).
- Lewis, J. S., Barshay, S. S. & Noyes, B. *Icarus* **37**, 190-206 (1979).
- Larimer, J. W. *Astrophys. Space Sci.* **65**, 351-369 (1979).
- Kurat, G. *Earth planet. Sci. Lett.* **7**, 317-324 (1970).
- Wood, J. A. *Geochim. cosmochim. Acta* **26**, 739-749 (1962).
- Müller, O. & Zähringer, J. *Earth planet. Sci. Lett.* **1**, 25-29 (1966).
- Rieder, R. & Wänke, H. in *Meteorite Research* (ed. Millman, P. M.) 75-86 (Reidel, Dordrecht, 1969).
- Bart, G. & Lipschutz, M. E. *Geochim. cosmochim. Acta* **43**, 1499-1504 (1979).
- Gaffey, M. J. & McCord, T. B. in *Asteroids* (ed. Gehrels, T.) 688-723 (1979).
- Johnson, T. V. & Fanale, F. P. *J. geophys. Res.* **85**, 8507-8518 (1973).
- Vdovykin, G. P. *Space Sci. Rev.* **14**, 758-831 (1973).
- JCPDS International Center for Diffraction Data, Card NO's 19-629 and 5-453 (1980).

Primary structure, gene organization and polypeptide expression of poliovirus RNA

Naomi Kitamura*, Bert L. Semler, Paul G. Rothberg, Glenn R. Larsen, Cheryl J. Adler, Andrew J. Dorner, Emilio A. Emini, Ronnie Hanecak, James J. Lee*, Sylvie van der Werf*, Carl W. Anderson† & Eckard Wimmer‡

Department of Microbiology, School of Medicine, State University of New York at Stony Brook, Stony Brook, New York 11794, USA and

† Department of Biology, Brookhaven National Laboratory, Upton, New York 11973, USA

* Present addresses: Department of Viral Oncology, Cancer Institute, 1-37-1 Kami-Ikebukuro, Toshima-ku, Tokyo 170, Japan (N.K.); Department of Biology, California Institute of Technology, Pasadena, California 91125, USA (J.J.L.); Unité de Virologie Moléculaire et Unité de Génie Génétique, Institut Pasteur, Paris, France (S.v.d.W.) ‡ To whom correspondence should be addressed

The primary structure of the poliovirus genome has been determined. The RNA molecule is 7,433 nucleotides long, polyadenylated at the 3' terminus, and covalently linked to a small protein (VPg) at the 5' terminus. An open reading frame of 2,207 consecutive triplets spans over 89% of the nucleotide sequence and codes for the viral polyprotein NCVPOO. Twelve viral polypeptides have been mapped by amino acid sequence analysis and were found to be proteolytic cleavage products of the polyprotein, cleavages occurring predominantly at Gln-Gly pairs.

POLIOVIRUS was discovered to be the causative agent of poliomyelitis by Landsteiner and Popper in 1908¹. Although development of both killed and live vaccines^{2,3} has been spectacularly successful in controlling poliomyelitis in industrialized countries since the mid 1950s, poliomyelitis remains a serious threat in much of the world. In spite of intensive research, the precise mechanisms by which poliovirus replicates and causes cytopathogenic effects, and the molecular basis for the attenuation of live virus vaccines, remain unknown.

We have now determined the primary nucleotide sequence of the entire single-stranded genome of poliovirus type 1 (Mahoney strain). Elucidation of the primary structure of the viral genome might reveal elements important to virus replication that might not have been apparent from other experimental approaches, such as the genome-linked protein VPg found attached to the 5' ribonucleotide⁴⁻⁶, the initiation and termination sites for poliovirus translation, structural signals in protein processing⁷ and potential elements of secondary structure within the genome itself⁸.

Genome RNA and viral mRNA are thought to be identical in nucleotide sequence and differ only in the presence or absence of VPg^{9,10}. Hence, studies of genome RNA permit an analysis of gene expression by translation. Poliovirus follows a unique strategy of gene expression by synthesizing a primary polypeptide which may contain all of the virus' translatable information¹¹ and which is cleaved post-translationally into numerous polypeptides¹¹⁻¹³. The relationships of these cleavage products to each other have been determined¹⁴. However, the signals directing protein processing, the proteinases involved, and the

possibility of regulating viral replication through polypeptide cleavage have only been matters of speculation.

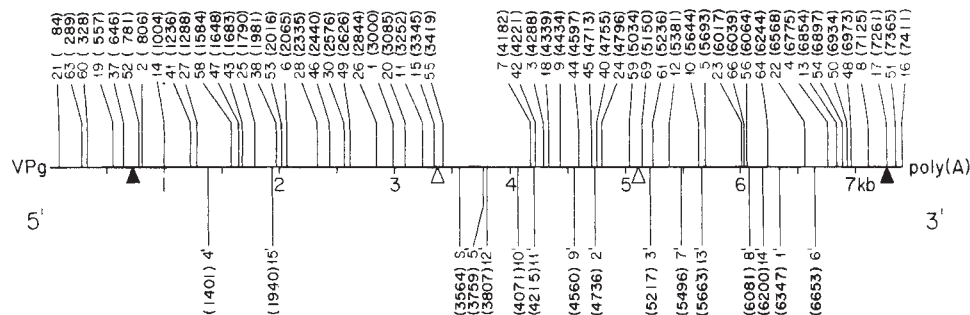
Sequence analysis of viral RNA

Poliovirus RNA was sequenced in three phases: (1) the identification of the 3'-terminal poly(A), (2) the determination of internal sequences, and (3) the characterization of a 5'-terminal, protein-linked fragment.

Following Kates¹⁶ discovery of polyadenylate in eukaryotic mRNA, poliovirus RNA was analysed and found to contain a poly(A) tract^{17,18} at its 3' end¹⁸. This structural element permitted the reverse transcription of poliovirus RNA into cDNA using oligo(dT) primers and reverse transcriptase^{19,20}.

Poliovirus RNA can be readily reverse-transcribed, so a logical strategy would be to clone the cDNA and then sequence the cloned DNA. Until recently this strategy of analysis was prohibited for poliovirus RNA by the NIH recombinant DNA guidelines. Instead, we sequenced poliovirus cDNA by a modification²⁰ of Sanger's chain termination method²¹ as follows: (1) Poliovirus cDNA was synthesized and chains of 7,000-7,400 deoxyribonucleotides were selected by zonal centrifugation. (2) Poliovirus RNA was digested exhaustively with RNase T₁ or with RNase A, and the large oligonucleotide products were separated by two-dimensional gel electrophoresis. Large oligonucleotides ($n > 10$) were eluted from the gel, dephosphorylated at their 3' ends and labelled at their 5' ends with phosphorus-32. (3) In four reaction mixtures the cDNA and a specific, 5'-³²P-labelled primer oligonucleotide were annealed and then incubated with *Escherichia coli*

Fig. 1 Physical map of RNase T₁- (upper row) and RNase A-resistant (lower row) oligonucleotides ($n > 10$) of poliovirus RNA. Numbers refer to oligonucleotides as identified by two-dimensional polyacrylamide gel electrophoresis^{20,24}. The location, within the RNA sequence, of the 5'-terminal nucleotide of an oligonucleotide is given in parentheses. With these locations the sequences of the oligonucleotides can be obtained from Fig. 2. With the exception of numbers 66, 56, 64, 54, 17, 51, 16, and 11', all oligonucleotides were used in the chain termination method. Note that oligonucleotides 19, 47 and 4' were poor primers, and 55, 15, 5' and 13' did not prime DNA synthesis at all. 'S' refers to a synthetic decaoxynucleotide (Collaborative Research) used as a primer to complete the sequence. ▲, Suggested sites of initiation and termination of translation; △, position of the primary cleavage sites of the polyprotein (see Fig. 4).



© 1981 Nature Publishing Group

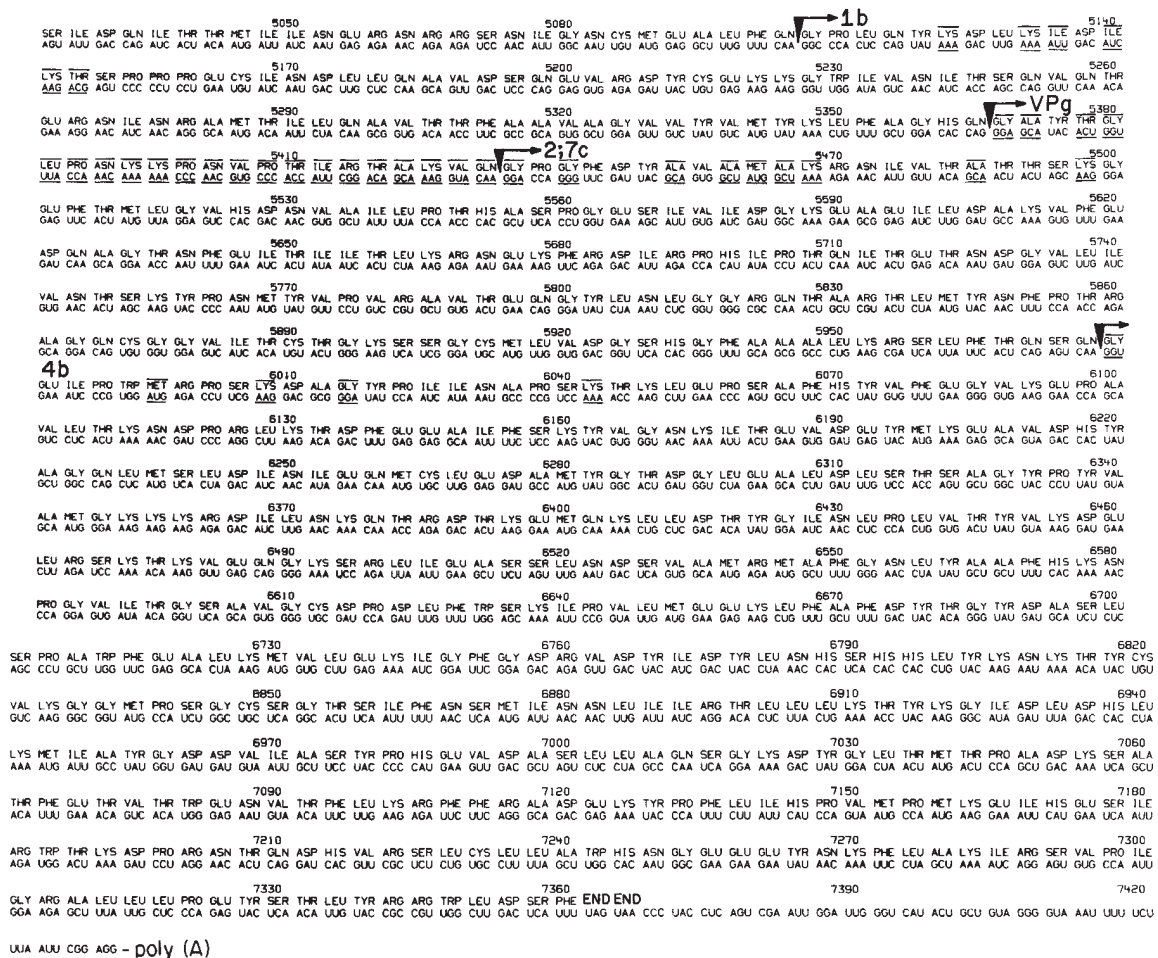


Fig. 2 The complete nucleotide sequence and encoded information of virion RNA of poliovirus type 1 (Mahoney). This strain was originally obtained from David Baltimore (Massachusetts Institute of Technology). As judged by fingerprint analyses this isolate has not undergone detectable genetic variation in our laboratory during multiple passages in HeLa cells²⁴. A reading frame of 2,207 consecutive coding triplets within the nucleotide sequence was determined by computer analysis and confirmed by radiochemical microsequence analysis of 12 virus-specific polypeptides. The cleavage sites involved in processing viral polypeptides and the N-termini of viral proteins are indicated by a solid triangle and arrow; amino acids that were identified by stepwise Edman degradation of radiochemically pure polypeptides are indicated by bars. The asterisk at N4159 indicates an amino acid ambiguity. The coding sequence of capsid protein VP4 has been determined by analysis of tryptic peptides. The N-terminal amino acid of VP4 is blocked as are those of P1-1a and VPO. Preliminary evidence suggests that translation may be initiated with the methionine at N741. Termination codons in the non-coding regions are designated as 'END'. Potential initiation codons of translation in the 5'-terminal, non-coding sequence preceding N741 are indicated as MET if in phase with the major reading frame. MET codons out of phase with the major reading frame in this region are indicated with brackets. 156 nucleotides preceding the 3'-terminal poly(A) have also been determined by Porter *et al.*⁶⁶. Analyses of the dinucleotide frequencies, codon usage and secondary structures will be published elsewhere (in preparation).

polymerase I (Klenow fragment) in the presence of unlabelled dNTPs and one of four 2',3'-dideoxynucleotide triphosphates (ddNTPs). (4) Synthesized DNA fragments (labelled through their 5'-terminal primer) were separated by slab gel electrophoresis²² and a sequence was determined by reading the four different 'ladders' of bands. We modified the chain termination procedure of Sanger only by using radiolabelled primers and unlabelled deoxynucleotide triphosphates. This modification reduced the background of radioactivity in gels to nearly undetectable levels.

In the early stages of the investigation the newly generated sequences were ordered according to the known map positions of the largest oligonucleotides²³. Subsequently, each newly generated sequence was searched in its 3'-terminal region for a sequence corresponding to another known, large T₁- or A-oligonucleotide²⁴. This oligonucleotide was then used as a primer to extend the sequence (for details, see ref. 20).

Figure 1 shows a physical map of large ($n > 10$) RNase T₁-resistant (upper row) and RNase A-resistant (lower row) oligonucleotides found in poliovirus RNA. With the exception of 'S', all of these oligonucleotides have been sequenced by rapid sequencing methods²⁴. Most of them were used as primers for

the chain termination method. With one exception (see below), each priming oligonucleotide produced a sequence of 250 bases, a segment long enough for overlaps with adjacent segments. Not all oligonucleotides primed DNA synthesis with equal efficiency²⁰, and four oligonucleotides (numbers 55, 15', 13' and 5'; ref. 24) did not prime DNA synthesis at all. Initiation at multiple sites with any single primer was not observed. 'Compressions' of bands were resolved either by slab gel electrophoresis at an elevated temperature²⁰ or by replacing dGTP in the reaction mixture with dITP²⁵ at an identical concentration, the latter being the method of choice.

The peculiar gap in T₁-oligonucleotides between nucleotides 3,419 and 4,182 (Fig. 1) may be without biological significance as this region is average in its G content. Unfortunately, the gap made it impossible to produce an overlap between N3570 and N3800. To solve this problem we obtained a custom-synthesized decadeoxyribonucleotide ['S', d(T-A-C-T-A-C-T-G-C-G), position N3564] that was selected on the basis of the known sequence generated by using T₁-oligonucleotide 15 as a primer for DNA synthesis. A reaction primed with 'S' did indeed produce a sequence that bridged this gap and yielded a continuous sequence of over 7,300 nucleotides. The 5'-terminal

sequence ($n \sim 100$), however, was still unknown because the yield of cDNA strands spanning the entire genome was too low for sequence analysis.

The study of the 5' end of poliovirus RNA proved difficult because no 5'-terminal nucleotide could be identified. Once the genome protein VPg was discovered and its covalent linkage to the 5'-terminal nucleotide established^{5,6}, it was possible to identify the 5'-terminal RNase T₁-resistant oligonucleotide (VPg-pU-U-A-A-A-A-C-A-G; refs 6,26). Moreover, Hewlett and Florkiewicz²⁷ succeeded in sequencing 20 nucleotides from the 5' end by partial digestion of ¹²⁵I-labelled VPg-RNA. A similar approach was developed by Nomoto and Imura²⁸ who obtained a sequence of ~ 30 nucleotides (A. Nomoto, personal communication).

The remaining gap in the 5'-terminal sequence was determined using several approaches. Poliovirus is cleaved by RNase III at low ionic strength into numerous fragments²⁹ of which two ($n = 85$ and $n = 145$) originate from the 5' end³⁰. The smaller RNase III fragment was sequenced by rapid techniques after 3'-labelling with [³²P]pCp, and by analysis of its T₁-resistant oligonucleotides⁸. The latter approach was necessary because of a stable stem-and-loop structure within the sequence that is resistant to digestion conditions of partial cleavage (see Fig. 3). The 5'-terminal sequence of the fragment was verified by analysis of poliovirus 5'-³²P-labelled mRNA⁸. Such analysis of the 5' end of mRNA was possible because poliovirus mRNA does not contain VPg but has a 5'-terminal pU³¹; mRNA can therefore be dephosphorylated and end-labelled with polynucleotide kinase and [γ -³²P]ATP. The sequences of the large RNase III fragments and of the 5' terminus of the mRNA connected with the sequence generated by the dideoxy method with T₁-oligonucleotide 21 (that maps at nucleotide 84) as the primer. This result completed the sequence analysis of poliovirus RNA.

The total nucleotide sequence of the heteropolymeric region of the genome of poliovirus type 1 (Mahoney) is shown in Fig. 2. Together with its 3'-terminal poly(A) (average chain length $n = 60$; see below), the genome RNA has a base composition of A = 30.2%, G = 22.8%, C = 23.1% and U = 23.9%, and a molecular weight of 2.411×10^6 (free acid).

We do not claim that the nucleotide sequence shown in Fig. 2 is entirely free of errors. One problem is that only one polynucleotide strand has been sequenced. Although single nucleotide assignments may require modification, the methods used make it unlikely that stretches of nucleotides have been overlooked. The following considerations lead us to accept the sequence shown in Fig. 2 as an accurate representation of the poliovirus genome: (1) The chain termination method, as we have applied it, was carried out to sequence a population of cDNA molecules rather than a single cloned DNA copy that may or may not be representative of genomic RNA. (2) All large oligonucleotides (16% of the total sequence) have been sequenced independently and by different methods²⁴. (3) A restriction enzyme map⁷⁵ of cloned cDNA agrees with that derived by computer analysis of our sequence. (4) The predicted sequence of NCVPOO is consistent with the partial amino acid sequence determined for 12 viral polypeptides.

VPg-linked 5'-terminal non-coding sequence

The 5' end of poliovirus RNA is highly unusual. Its nucleotide sequence up to the first RNase III cleavage site is shown in Fig. 3. The sequence contains a stem-and-loop structure, corresponding to nucleotides 9–36 that could be involved in processes of regulation or recognition⁸. The 5'-terminal uridine is linked, via a phosphodiester bond, to a tyrosine residue of the genome protein, VPg^{5,6,32–34}. VPg is a virus-encoded^{5,35,36}, basic ($pI > 9.4$; unpublished results, and ref. 6), acid-soluble, water-insoluble oligopeptide consisting of only 22 amino acids^{7,36} (Fig. 2). It maps in P3, a region of the genome RNA specifying

replication proteins^{7,36,37}. The O^4 -(5'-uridylyl) tyrosine bond is energy rich³⁸, but the mechanism by which it is formed is obscure, as is the function of VPg. Since poliovirus minus strand RNA has VPg linked to the 5'-terminal poly(U)^{6,39,40} and because nascent RNA strands of poliovirus replicative intermediates are VPg linked^{6,40}, VPg has been proposed to have a role in the initiation of RNA synthesis. In addition, or alternatively, VPg may be involved in morphogenesis (reviewed in refs 41, 42).

The data suggest that translation of the large polypeptide begins within the large open reading frame at the AUG codon located at N741. This leaves a 5' sequence of 740 potentially untranslated nucleotides. Preceding the left-most AUG in the long, open reading frame at N741 are several other AUG codons. Specifically, these occur at N185, N237, N276, N318, N376, N387, N455, and N584 (those underlined being in phase with the major reading frame). All but three of these AUG codons are followed closely by stop codons within their reading frame. However, three of the AUG triplets (at N185, N376 and N584, underscored by dotted lines) are not in phase with the polypeptide but they could initiate synthesis of polypeptides of molecular weights 4,725, 8,425, and 7,187, respectively. The significance of these potential coding sequences is not clear. The following explanations are possible: (1) Sequencing errors occurred between N400 and N774 that disconnected the shorter reading frames from the major reading frame. This explanation is unlikely because the sequence between 400 and 793 has been corroborated with cloned cDNA of poliovirus of the same strain

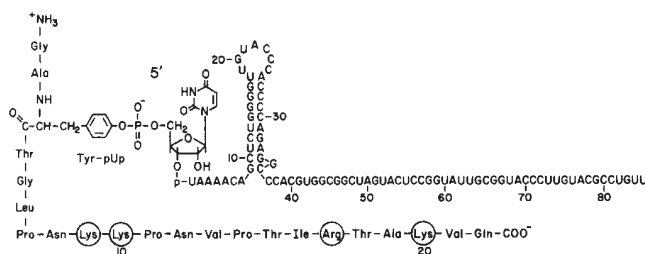


Fig. 3 Structure of the 5'-terminus of poliovirus RNA. The oligonucleotide ($n = 85$) is an RNase III cleavage fragment of virion RNA⁸. The viral protein VPg is linked via an O^4 -(5'-uridylyl)tyrosine (Tyr-pUp) bond to the RNA. The phosphodiester bond in Tyr-pUp^{32,33} can be considered energy rich by analogy to O^4 -(5'-adenylyl)tyrosine³⁸. The stem-and-loop structure as drawn here has a ΔG° of $-21 \text{ kcal mol}^{-1}$. The nature of the interaction between the basic genome protein (see encircled amino acids) and the RNA is unknown.

(S.v.d.W., E. W. and M. Girard, unpublished results). (2) Poliovirus codes for one or more small proteins independently from the large polypeptide. Such small proteins may have missed detection because of their size, their lack of internal methionines, or because of rapid degradation. Ehrenfeld and co-workers⁴³ have detected a small polypeptide when poliovirus RNA is translated *in vitro* in the presence of 4 mM MgCl₂. It remains to be seen whether this product is encoded by one of the above potential coding segments. (3) The small coding segments remain untranslated.

Identification of the AUG codon that functions in initiating translation of the polypeptide has not yet been possible. Assignment of the correct initiation codon is complicated by the following observations. (1) Ribosome binding experiments as carried out according to Steitz⁴⁴ have failed in this and other laboratories (E. Ehrenfeld, personal communication, and ref. 45). The major obstacle was that the ribosomes were found to bind to more than one site. Digestion of the ribosome-bound RNA with RNase yielded RNA fragments of equal length but different sequence (A. J. D. and E. W., unpublished results). (2) An equivalent to the sequence found in prokaryotic mRNA,

that directs the ribosome to the initiation codon of translation⁴⁶, does not seem to exist in eukaryotic mRNA⁴⁷. (3) The polypeptides derived from the N-terminus of the polyprotein (P1-la, VPO and VP4) are blocked at their N-termini (unpublished results). Regardless of whether the small reading frames are translated *in vivo*, the occurrence of several AUG codons in different reading frames within this region makes it likely that some AUG codons are ignored in viral translation.

Fine structure of the protein map

In the nucleic acid sequence of poliovirus RNA (Fig. 2) an open reading frame of 2,207 consecutive triplets spans over 89% of the entire viral genome and has the coding capacity for a polypeptide of 247,000 *M_r*. The length of the reading frame is strong evidence that it codes for the polyprotein NCVPOO, the polypeptide proposed to represent the entire translatable information of poliovirus¹¹. This conclusion is further supported by the following observations: (1) No other open reading frame longer than 78 triplets exists within the sequence. (2) All major viral polypeptides, 12 of which we have analysed by protein sequencing, map in phase within the long reading frame. These data prove that all known poliovirus proteins can originate by proteolytic cleavage from a single precursor polypeptide¹¹⁻¹³.

Individual poliovirus proteins were positioned with respect to the nucleotide sequence by partial amino acid sequence analysis and by alignment of the predicted with the experimental sequences. Viral polypeptides were labelled *in vivo* with one amino acid at a time, purified by polyacrylamide gel electrophoresis and their N-terminal amino acid sequences determined by radiochemical microsequence analysis^{7,36}. A computer search of the poliovirus genomic coding sequence identified for each amino acid pattern a unique encoding location (see Fig. 2).

We have divided the polypeptide map into three regions (see Fig. 4): P1 corresponds to the region of capsid polypeptides and their precursors, P2 corresponds to the central portion of the genome, and P3 to a group of non-structural polypeptides including the putative proteinase P3-7c (by analogy with studies on encephalomyocarditis virus)^{48,49}, VPg^{7,36} and the RNA-dependent RNA polymerase P3-4b^{7,50,51}. Our data unambiguously show that, contrary to a previous proposal¹⁵, P2-X maps within the centre of the viral genome.

Our strategy of locating the N-terminus of a protein with respect to the genomic sequence failed for components P1-la, VPO and VP4 because these polypeptides have blocked N-termini. Nonetheless, we could deduce limits for the sequence that encodes VP4 from a knowledge of its estimated molecular weight (corresponding to ~70 amino acids), its reported amino acid composition^{52,53}, and the fact that it is the only known protein that maps to the left of VP2 (reviewed in ref. 14). Within 70 amino acids to the left of VP2, only one tyrosine-containing tryptic peptide is predicted at N794-N842 (with Tyr at positions 2, 9, 14 and 15), and that was confirmed experimentally by stepwise microsequence analysis of a total trypsin digest of ³H-Tyr-labelled VP4 (see Fig. 2; unpublished results). The N-terminal boundary of VP4 must be to the right of N734, as a cysteine is predicted at N732 and VP4 does not contain cysteine^{52,53}. Three codons to the right of the cysteine (N732) the genomic sequence predicts a methionine codon (AUG at N741) which is the left-most methionine codon within the large open reading frame shown in Fig. 2. This AUG codon is a likely candidate for the initiation codon of the precursor polypeptides to VP4 (see below).

The next AUG codon within the open reading frame occurs at N939, only three codons to the left of the N-terminus of VP2. This methionine would be contained within the C-terminal tryptic peptide of VP4. Analysis of a tryptic digest of ³⁵S-Met-labelled VP4 by HPLC yielded a single Met-containing tryptic peptide. Stepwise degradation showed a Met in position 4 as would be expected for the predicted C-terminal tryptic peptide

Table 1 Features of poliovirus proteins

Protein	Position in nucleotide sequence	No. of amino acids	Molecular weights	Net charge
NCVPOO	741-7,361	2,207	246,530	+44
P1-1a	744-3,380	879	97,247	+12
VPO	744-1,760	339	37,352	+3
VP4	744-947	68	7,385	+2
VP2	948-1,760	271	29,985	+1
VP3	1,761-2,474	238	26,410	0
VP1	2,475-3,380	302	33,521	+9
P2-3b	3,381-5,105	575	64,953	+16
P2-5b	3,828-5,105	426	48,273	+15
P2-X	4,119-5,105	329	37,555	+15
P3-1b	5,106-7,361	752	84,234	+16
VPg	5,367-5,432	22	2,354	+4
P3-2	5,433-7,361	643	72,132	+11
P3-4b	5,979-7,361	461	52,481	+6
P3-7c	5,433-5,978	182	19,669	+5

The initiation site of translation at N741 (NCVPOO) is based on circumstantial evidence only. The number of amino acids and molecular weights are given on the assumption that C-terminal trimming of polypeptides does not occur.

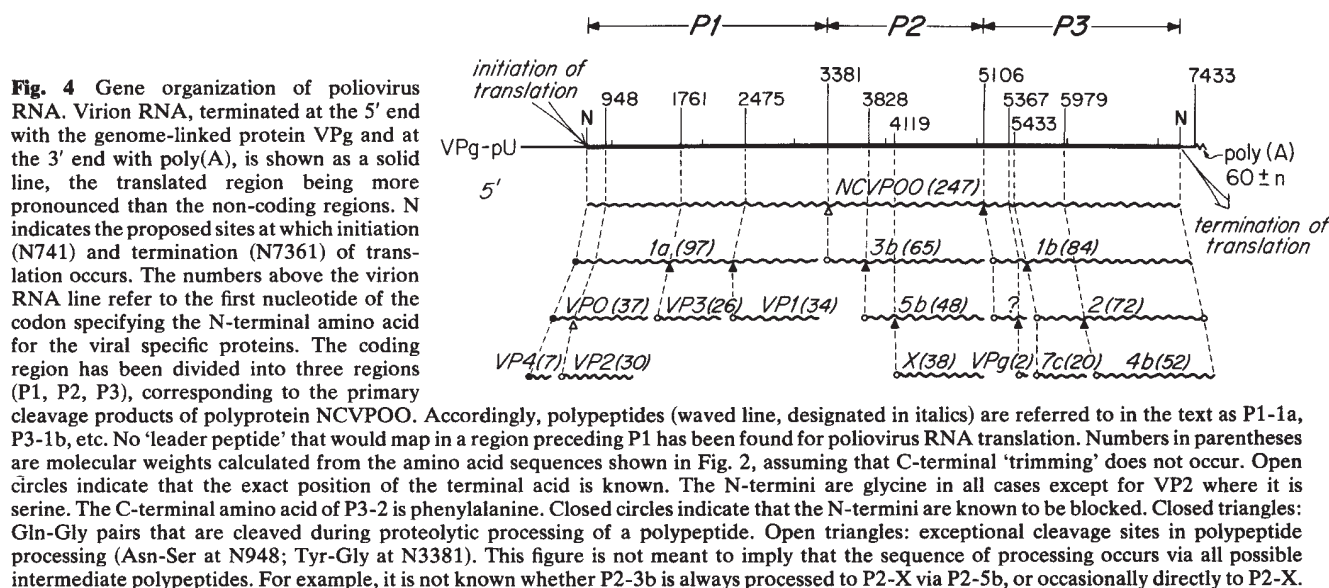
(unpublished results). The presence of a single Met-containing tryptic peptide in VP4 is consistent with the removal of the methionine at N741 by analogy to previously sequenced proteins⁵⁴. We predict that the N-terminus of VP4 arises from an expected modification of the nascent polypeptide and that it has the sequence, beginning at N744, of (block)-Gly-Ala-Gln-Val-Ser-Ser-Gln-Lys-. VP4 would then contain 68 amino acids and have a predicted molecular weight of ~7,500.

Cell-free translation experiments with aphthovirus RNA have demonstrated the existence of a polypeptide (16,000 *M_r*) that maps in a region preceding the P1 region⁵⁵. The results of *in vitro* translation of encephalomyocarditis (EMC) virus RNA have also been interpreted to mean that a 'leader' peptide precedes the capsid polypeptides, but direct evidence for such a leader is lacking (for a discussion, see ref. 43). On the other hand, translation of poliovirus RNA *in vitro* in the presence of labelled formyl-methionyl-tRNA has failed to produce evidence for the existence of a comparable poliovirus 'leader peptide' (E. Ehrenfeld, personal communication; our unpublished results).

The molecular weights (Table 1) calculated from the predicted amino acid sequences (Fig. 2) agree well with those deduced from SDS-polyacrylamide gel electrophoresis. The calculations are based on the assumption that no trimming of the C-termini occurs after processing. Moreover, the sequences of capsid proteins shown in Fig. 2 agree well with published data on N-terminal amino acid sequences and amino acid compositions of poliovirus type 1 (strain 1a/S3) capsid proteins^{52,56}, and with the amino acid composition of type 1 (Mahoney) capsid polypeptides⁵⁷. Heterogeneity of the capsid polypeptides, as reported by Vrijnsen *et al.*⁵⁶, has not been observed in our studies.

The C-terminal amino acid of polypeptide P3-2 was determined by carboxypeptidase analysis to be phenylalanine (unpublished results). This observation and the results of nucleotide sequence analyses suggest that translation of poliovirus terminates at N7361, 73 nucleotides before the beginning of the poly(A) tract. (We previously suggested that termination of translation may occur at N6874^{20,42}. This result was erroneous due to three sequence ambiguities.)

The genetic map that we have derived from our protein and RNA sequence analyses is summarized in Fig. 4. It is in agreement with a map published by Rueckert and his colleagues⁵⁸. Our data extend this information by indicating the precise location of the N-terminus of each major poliovirus proteolytic cleavage product with respect to the genomic sequence. Products of alternate cleavage pathways are not



shown but will be discussed in a later section. Omitted are also several small virus-specific peptides ($M_r < 12,000$) that can be observed in PAGE analyses of infected cells, whose origins and functions, however, are unknown.

The 3'-terminal, non-coding sequence

Our data suggest that the 3'-terminal, non-coding sequence of poliovirus RNA consists of a heteropolymeric segment of 72 nucleotides followed by poly(A). It lacks the polyadenylation signal A-A-U-A-A-A- found in polyadenylated mRNAs of eukaryotic cells⁵⁹. Lack of this hexanucleotide supports our notion that the poly(A) of poliovirus RNA is genetically coded in that it is transcribed from 5'-terminal poly(U) in viral minus-stranded RNA^{39,60-62}. Poly(A) in poliovirus RNA is very heterogeneous in length¹⁸, but the average chain length has been determined directly by rapid sequencing techniques to be 60 residues long⁶³. The function of poly(A) in poliovirus RNA is obscure, but it may have a role in the initiation of synthesis of minus-stranded RNA^{64,65}.

Signals in proteolytic processing

With a single exception (VP2) all viral polypeptides that we have analysed have glycine as their N-terminal amino acid (Fig. 2; refs 7, 36). Inspection of the sequences of the precursor polypeptides reveals that in 8 of 9 cases the NH₂-terminal Gly of the cleavage products was preceded by a glutamine residue. These data strongly suggest that proteolytic processing of poliovirus polypeptides occurs predominantly at Gln-Gly pairs in all three regions of the polyprotein, and that these cleavages are carried out by the same proteinase.

It has been assumed that host proteinase(s) act on the polyprotein NCVPOO to cut it *in statu nascendi* into the 'primary' cleavage products P1-1a, P2-X and P3-1b⁶⁷. Proteolytic processing of P1-1a and P3-1b into 'secondary' cleavage products, on the other hand, has been clearly shown to be catalysed by virus-coded enzyme(s)^{48,49,68}. We now consider it likely that the primary cleavage that generates P3-1b and secondary cleavages are carried out by the viral proteinase(s) because (1) the cleavage sites recognized in these processes are identical or similar (Fig. 2), and (2) a viral polypeptide can undergo autocatalytic cleavage to yield a viral proteinase and a putative RNA polymerase. The latter case exists in EMC virus, where polypeptide D (corresponding to P3-2 in polio) cleaves to proteinase p22 and E⁶⁹, corresponding to P3-7c and P3-4b,

respectively, in poliovirus. Assuming that an analogous situation exists for poliovirus, it is tempting to speculate that during initial stages of infection primary cleavages result from autocatalytic events or by *trans* action of one polypeptide on another. As the infection proceeds, sufficient amounts of mature viral proteinase are produced to catalyse subsequently primary as well as secondary cleavages.

Inspection of the amino acid sequence illustrated in Fig. 2 shows the dipeptide Gln-Gly to occur 13 times in the polyprotein NCVPOO, 2.5 times the expected number of occurrences for this amino acid pair calculated assuming random distribution of amino acids of the composition and size of the polyprotein. The high Gln-Gly frequency supports the hypothesis that Gln-Gly is a major signal for processing of poliovirus proteins. At least 8 of the 13 Gln-Gly pairs are utilized in processing. Gln-Gly cleavage may also occur at N5796 to yield polypeptide P3-4a⁷⁰, and at N6486 to yield the products P3-6a and P3-6b of an alternate cleavage pathway within P3-1b⁵⁸. Amino acid sequence analyses of these products are now in progress. Remaining Gln-Gly pairs at N1302, N3519 and N4626 may be ignored in processing either because the amino acid sequences surrounding these dipeptides may protect them from the proteinase, or because these cleavage signals may be sequestered. Alternatively, these Gln-Gly pairs may be utilized to a lesser degree so that the corresponding cleavage products have not been detected. Preliminary studies of the amino acid sequences around the Gln-Gly pairs that are known processing sites have not revealed other constant features likely to contribute to processing specificity. Other processing determinants besides the Gln-Gly sequence must exist, however, because infection with poliovirus does not result in detectable degradation of host cellular proteins, at least not for the first 4 h during which the bulk of viral macromolecular synthesis and proteolytic processing occurs⁶⁷.

One amino acid pair of interest is the Asn-Ser in VPO (at N945) that is cleaved during maturation of the virion to yield VP4 and VP2 (reviewed in ref. 14). Asn-Ser (two helix-breaker residues) differs from Gln-Gly (helix maker-helix breaker residues)⁷¹ and could therefore be part of a signal for a processing step at maturation catalysed by a unique proteinase. A possible candidate, P2-X, has been reported to have proteolytic activity⁷². Note that the Asn-Ser dipeptide is predicted to occur at seven other locations in the polyprotein, all but two of which are within the capsid region.

The cleavage that generates the N-terminus of P2-3b is predicted by our sequence (Fig. 2) to occur between a tyrosine and a glycine residue. We have confirmed that the N-terminus of P2-3b is glycine (unpublished results); however, the identity of

the C-terminus of VP1 has not yet been directly determined. Thus we cannot yet be certain whether this site is unique and is cleaved by another, perhaps cellular, proteinase, whether it is cleaved by the activity that recognizes Gln-Gly pairs, or whether a sequencing error has resulted in the erroneous prediction of tyrosine for this position.

Previous studies have shown that a Gln-Gly may be cleaved between the VP3 and VP1 of mengoviruses⁷³. In contrast to our data, the NH₂-terminal amino acids of mengovirus are Asp in VP2, Ser in VP3, and Gly in VP1⁷³, and of aphthovirus are Asp in VP2, Gly in VP3, and Thr in VP1⁷⁶. Burrell and Cooper⁷⁴ reported Gly, Asp and Ser as N-terminal amino acids in a mixture of poliovirus capsid proteins. Our nucleotide sequence predicts at the N-termini Ser in VP2 and Gly in VP3 and VP1, which we have confirmed by radiochemical microsequence analysis (Fig. 2). Thus, a discrepancy exists that cannot be explained at the present time.

Conclusions

To our knowledge, this work is the first complete description of the primary structure of a eukaryotic, helper-independent RNA virus genome. The protein-linked genome represents a unique chemical entity not previously found among naturally occurring RNAs.

The nucleotide sequence of the viral RNA is useful in several ways. In combination with radiochemical microsequence analyses we have determined the coding regions for each major viral polypeptide and thus their amino acid sequences. Based on the known virion structure¹⁴ and the information presented in this article, the chemical structure of the poliovirion (calculated molecular weight 8.25×10^6) has been elucidated. Knowledge of

the amino acid sequence will aid in the isolation and characterization of the viral gene products, for example, by antisera elicited by specific synthetic oligopeptides. The sequence data have also been instrumental in detecting and mapping cloned fragments of poliovirus cDNA. Specifically, the data have been useful in the prediction of cleavage sites of restriction enzymes and in the screening of clones by hybridization of 5'-³²P-labelled oligonucleotides (Fig. 1) to the DNA in lysates of plasmid-containing *E. coli*⁷⁵.

The development of new vaccines by novel methods is desirable, and one approach is the construction of attenuated viral strains incapable of reverting to virulence²⁴. A second approach could make use of the expression of cloned virus-specific genome segments in *E. coli* to produce viral antigens that are capable of eliciting formation of poliovirus neutralizing antibodies in man. Pursuit of both approaches requires that the genome structure and the gene organization of poliovirus be known.

Finally, construction of specific inhibitors of the viral proteinase(s) utilizing a knowledge of their cleavage specificity may lead to a novel set of antiviral chemotherapeutic agents⁶⁷.

We thank Akiko Kitamura, Lydia Fedus and Medora Hardy for technical assistance. We are indebted to Astrid Wimmer for support, to Carol Carter, Ann Jacobson, Nancy Reich and Akio Nomoto for helpful discussions, and to Keith Thompson for assistance in computing data. We also thank Joyce Schirmer and Jeffrey Demian for art work. Computer services were provided by the Stanford MOLGEN project of the NIH SUMEX-AIM facility. This work was supported in part by PHS grants AI 15122 and CA 28146; and by the US Department of Energy. B.L.S. is supported by NIH fellowship AI 05935; E.A.E. is a postdoctoral trainee; and C.J.A. a predoctoral trainee of NCI, grant 09176. E.W. is a Josiah Macy Jr Faculty Scholar.

Received 16 February; accepted 8 April 1981.

- Landsteiner, K. & Popper, E. *Wien. Klin. Wochenschr.* **21**, 1830 (1908).
- Salk, J. & Salk, D. *Science* **195**, 834-847 (1977).
- Sabin, A. B. & Boulger, C. R. *J. biol. Standard* **1**, 115-118 (1973).
- Lee, Y. F., Nomoto, A. & Wimmer, E. *Prog. Nucleic Acid Res. molec. Biol.* **19**, 89-95 (1976).
- Lee, Y. F., Nomoto, A., Detjen, B. M. & Wimmer, E. *Proc. natn. Acad. Sci. U.S.A.* **74**, 59-63 (1977).
- Nomoto, A., Detjen, B., Pozzatti, R. & Wimmer, E. *Nature* **268**, 208-213 (1977).
- Semler, B. L. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Larsen, G. R., Semler, B. L. & Wimmer, E. *J. Virol.* **37**, 328-335 (1981).
- Pettersson, R. F., Flanagan, J. B., Rose, J. K. & Baltimore, D. *Nature* **268**, 270-272 (1977).
- Nomoto, A., Kitamura, N., Golini, F. & Wimmer, E. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5345-5349 (1977).
- Jacobson, M. F. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **61**, 77-84 (1968).
- Summers, D. F. & Maizel, J. V. *Proc. natn. Acad. Sci. U.S.A.* **59**, 966-971 (1968).
- Holland, J. J. & Kiehn, E. D. *Proc. natn. Acad. Sci. U.S.A.* **60**, 1015-1022 (1968).
- Rueckert, R. R. *Compreh. Virol.* **6**, 131-219 (1976).
- Cooper, P. D. *Compreh. Virol.* **9**, 133-207 (1977).
- Kates, J. *Cold Spring Harb. Symp. quant. Biol.* **35**, 743-752 (1970).
- Armstrong, J. A., Edmonds, M., Nakazato, H., Philips, B. A. & Vaughan, M. H. *Science* **176**, 526-528 (1972).
- Yogo, Y. & Wimmer, E. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1877-1882 (1972).
- Kacian, D. L. & Myers, J. C. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2191-2195 (1976).
- Kitamura, N. & Wimmer, E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3196-3200 (1980).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Sanger, F. & Coulson, A. R. *FEBS Lett.* **87**, 107-116 (1978).
- Nomoto, A., Jacobson, A., Lee, Y. F., Dunn, J. J. & Wimmer, E. *J. molec. Biol.* **128**, 179-196 (1979).
- Nomoto, A. *et al. Virology* (in the press).
- Kramer, F. R. & Mills, D. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5334-5338 (1978).
- Flanagan, J. B., Pettersson, R. F., Ambros, V., Hewlett, J. J. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 961-965 (1977).
- Hewlett, M. J. & Florkiewicz, R. Z. *Proc. natn. Acad. Sci. U.S.A.* **77**, 303-307 (1980).
- Nomoto, A. & Imura, N. *Nucleic Acids Res.* **7**, 1233-1246 (1979).
- Nomoto, A. *et al. J. molec. Biol.* **128**, 165-177 (1979).
- Harris, T. J. R., Dunn, J. J. & Wimmer, E. *Nucleic Acids Res.* **5**, 4039-4054 (1978).
- Nomoto, A., Lee, Y. F. & Wimmer, E. *Proc. natn. Acad. Sci. U.S.A.* **73**, 375-380 (1976).
- Rothberg, P. G., Harris, T. J. R., Nomoto, A. & Wimmer, E. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4868-4872 (1978).
- Ambros, V. & Baltimore, D. *J. biol. Chem.* **253**, 5263-5266 (1978).
- Wu, M., Davidson, N. & Wimmer, E. *Nucleic Acids Res.* **5**, 4711-4723 (1978).
- Golini, F., Nomoto, A. & Wimmer, E. *Virology* **89**, 112-118 (1978).
- Kitamura, N. *et al. Cell* **21**, 295-302 (1980).
- Pallansch, M. A. *et al. J. Virol.* **35**, 414-419 (1980).
- Holzer, H. & Wohlhueter, R. *Adv. Enzyme Reg.* **10**, 121-132 (1972).
- Yogo, Y. & Wimmer, E. *J. molec. Biol.* **92**, 467-477 (1975).
- Pettersson, R. F., Ambros, V. & Baltimore, D. *J. Virol.* **27**, 357-365 (1978).
- Wimmer, E. in *The Molecular Biology of Picornaviruses* (ed. Pérez-Bercoff, R.) 175-190 (Plenum, New York, 1979).
- Kitamura, N., Adler, C. & Wimmer, E. *Ann. N.Y. Academy Sci.* **354**, 183-201 (1980).
- Ehrenfeld, E. in *The Molecular Biology of Picornaviruses* (ed. Pérez-Bercoff, R.) 223-238 (Plenum, New York, 1979).
- Steitz, J. A. in *RNA Phages* (ed. Zinder, N. D.) 319-352 (Cold Spring Harbor Laboratory, New York, 1975).
- Golini, F., Semler, B. L., Dorner, A. J. & Wimmer, E. *Nature* **287**, 600-603 (1980).
- Shine, J. & Dalgarno, L. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1342-1346 (1974).
- Kozak, M. in *Protein Biosynthesis in Eukaryotes* (Plenum, New York, in the press).
- Palmenberg, A. C., Pallansch, M. A. & Rueckert, R. R. *J. Virol.* **32**, 770-778 (1979).
- Svitkin, Yu. V., Gorbalenya, A. E., Kazachkov, Yu. A. & Agol, V. I. *FEBS Lett.* **108**, 6-9 (1979).
- Lundquist, R. E., Ehrenfeld, E. & Maizel, J. V. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4773-4777 (1974).
- Flanagan, J. B. & Baltimore, D. *J. Virol.* **29**, 352-360 (1979).
- Wouters, M. & Vandekerckhove, J. J. *gen. Virol.* **33**, 529-533 (1976).
- Wetz, K. & Habermehl, K. O. *J. gen. Virol.* **44**, 525-534 (1979).
- Anderson, C. W. & Lewis, J. B. *Virology* **104**, 27-41 (1980).
- Sanger, D. V., Black, D. N., Rowlands, D. J., Harris, T. J. R. & Brown, F. J. *J. Virol.* **33**, 59-68 (1980).
- Vrijns, R., Boëye, A. & Strosberg, A. D. *Biochem. biophys. Res. Commun.* **85**, 1596-1601 (1978).
- Heukeshoven, J. & Dernick, R. Z. *Naturforschung* **36**, 164-172 (1981).
- Rueckert, R. R., Matthews, T. J., Kew, O. M., Pallansch, M., McLean, C. & Omilianowski, D. in *The Molecular Biology of Picornaviruses* (ed. Pérez-Bercoff, R.) 113-125 (Plenum, New York, 1979).
- Fitzgerald, M. & Shenk, T. *Cell* **24**, 251-260 (1981).
- Yogo, Y., Teng, M. H. & Wimmer, E. *Biochem. biophys. Res. Commun.* **61**, 1101-1109 (1974).
- Spector, D. H. & Baltimore, D. *Virology* **67**, 498-505 (1975).
- Dorsch-Häslar, K., Yogo, Y. & Wimmer, E. *J. Virol.* **16**, 1512-1527 (1975).
- Ahlquist, P. & Kaesberg, P. *Nucleic Acids Res.* **7**, 1195-1204 (1979).
- Spector, D. H. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **71**, 2983-2987 (1974).
- Dasgupta, A., Zabel, P. & Baltimore, D. *Cell* **19**, 423-429 (1980).
- Porter, A. G. *et al. Nature* **276**, 298-301 (1978).
- Korant, B. D., Chow, N., Lively, M. & Powers, J. *Ann. N.Y. Acad. Sci.* **343**, 304-315 (1980).
- Pelham, H. R. B. *Eur. J. Biochem.* **85**, 457-462 (1978).
- Rueckert, R. R., Palmenberg, A. C. & Pallansch, M. A. *Eur. J. cell. Biol.* **22**, 263-275 (1980).
- Etchison, D. & Ehrenfeld, E. *Virology* **107**, 135-143 (1980).
- Anfinson, C. B. & Scheraga, H. H. *Adv. Protein Chem.* **25**, 205-300 (1975).
- Korant, B., Chow, N., Lively, M. & Powers, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2992-2995 (1979).
- Scrab, D. G. in *The Molecular Biology of Picornaviruses* (ed. Pérez-Bercoff, R.) 1-23 (Plenum, New York, 1979).
- Burrell, C. J. & Cooper, P. D. *J. gen. Virol.* **21**, 443-451 (1973).
- Van der Werf, S. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Boothroyd, J. C. *et al. Nature* **290**, 800-802 (1981).

Effects of sleep and arousal on the processing of visual information in the cat

Margaret S. Livingstone* & David H. Hubel

Department of Neurobiology, Harvard Medical School, Boston, Massachusetts 02115, USA

Single units in the cat lateral geniculate nucleus and primary visual cortex show changes in both spontaneous and visually evoked firing as a function of the state of wakefulness. On arousal spontaneous firing is smoother and often reduced, whereas evoked responses are usually enhanced; the result is an increase in the signal-to-noise ratio. Single- and double-label 2-deoxyglucose autoradiographs show further that slow-wave sleep differentially depresses visually evoked activity in the deeper layers of the visual cortex.

SUBJECTIVELY it seems obvious that the way in which the brain handles sensory information must change radically between sleeping and waking. To examine these changes it is important to be able to influence the firing of neurones along a sensory pathway by natural stimulation. We have chosen the visual system to study neural activity as a function of the state of arousal because, in the pathway from the retina through the lateral geniculate body (LGB) to the primary visual cortex, the properties of the cells are well understood in terms of their optimum visual stimuli. In an animal that closes its eyes when asleep, additional mechanisms for blocking visual messages during sleep might seem unnecessary but, beyond the retina, such mechanisms do seem to exist. In the cat, spontaneous firing patterns of cells in the LGB^{1,2} and primary visual cortex³⁻⁶ show marked differences between sleeping and waking. In the LGB Coenen and Vendrik⁷ have shown that responses to light stimulation are strongly enhanced on arousal in lightly anaesthetized cats. In the present study we have recorded from single cells in the LGB and striate cortex⁸, examining the effects of the state of wakefulness on both the vigour and specificity of responses. We have also compared sleep and arousal using the 2-deoxyglucose method for labelling neurones according to activity^{9,10}.

Single-cell recording

We compared spontaneous and visually evoked firing in sleeping and waking states in 130 cortical cells in 15 cats, and in 14 geniculate neurones in 2 cats.

Each cat was first sleep deprived overnight in a slowly revolving (0.5 r.p.m.) drum, and then prepared as if for chronic recording^{1,4}, under halothane anaesthesia. Four chlorided silver electroencephalogram (EEG) electrodes were cemented over the dura and a pedestal for holding a hydraulic microelectrode advancer was cemented to the skull. For cortical recording we positioned the pedestal over the postlateral gyrus. For geniculate recording the animal's head was put in a Horsley-Clarke headholder and the pedestal cemented stereotaxically above the LGB. The head was then freed, well before terminating the anaesthesia. We infiltrated skin incisions with Xylocaine, intubated the trachea, paralysed the animal with continuous intravenous succinylcholine and monitored levels of expired CO₂. The head was held firmly with adhesive tape and gauze pads. Pupils were dilated with homatropine, contact lenses were fitted so that images on a screen at 1.5 m distance produced a focused image on the retina, and the optic disk and area centralis of each eye were mapped by optical projection. When necessary we used an adjustable prism to bring the two centres of gaze into superposition.

The halothane was then discontinued, and in the EEG irregular high-voltage activity of the anaesthetized state gave way, in a few minutes, to the low-voltage rapid activity of the waking state. Most of the cats then slept spontaneously and frequently, as indicated by high-voltage slow waves. Arousal from slow-wave sleep was spontaneous or could be produced by a noise or tactile stimulus. Rapid eye movement (REM) sleep was recognized by episodes of large monophasic potentials recorded from the visual cortex (ponto-geniculo-occipital (PGO) waves), which could be suppressed by arousing the animal.

For each cell we first mapped receptive fields on the screen with a hand-held slide projector. For cortical cells a computer-driven optic bench was adjusted to produce moving-slit stimuli which were tailored for optimal position, orientation, shape and rate of movement. The stimuli for geniculate cells were stationary circular spots or annuli. Action potentials, EEGs and traces representing the presentations of the stimuli were recorded on a Grass polygraph and on magnetic tape for subsequent averaging. Every cell included in the series was followed over several transitions between waking and high-voltage slow-wave sleep.

Recordings from lateral geniculate body

With the screen diffusely lit, geniculate cells showed an increase in spontaneous firing rate when the animal was aroused from slow-wave sleep, as expected from previous studies^{7,11-13}. Moreover, the pattern of firing changed from high-frequency bursts of two to eight spikes to more regular firing^{1,3,14}.

In all 14 geniculate cells there were marked changes also in responsiveness to visual stimuli, as illustrated in Fig. 1. In Fig. 1a, during an otherwise wakeful period, there was a brief episode of high-voltage slow-wave sleep lasting about 30 s and ending spontaneously. The stimulus, which was repeated throughout, was a small (1/4°) spot of light covering the receptive field centre. During the slow waves there was a depression of the on-response, obvious in both the spike record and the average-response histograms. This change is most easily explained by supposing that the efficacy of optic nerve fibres in triggering the geniculate cell is enhanced on arousal⁷. All the effects of arousal cannot be explained in this way, however. In Fig. 1b, the stimulus was an annulus covering the surround of the cell's receptive field. Now spontaneous arousal was accompanied not only by an increase in the off discharge, but also by an enhanced suppression of firing while the annulus was on.

Figure 1c shows the enhancement of suppression for the same on-centre cell, in a slightly different way. In Fig. 1c, upper half, the screen was lit steadily and diffusely by a low-level background light. The tonic firing of the cell was obviously decreased during each of the two bursts of slow waves. In the lower half, an annulus covering the receptive field surround tonically sup-

* Present address: Department of Biology, Princeton University, Princeton, New Jersey 08544, USA.

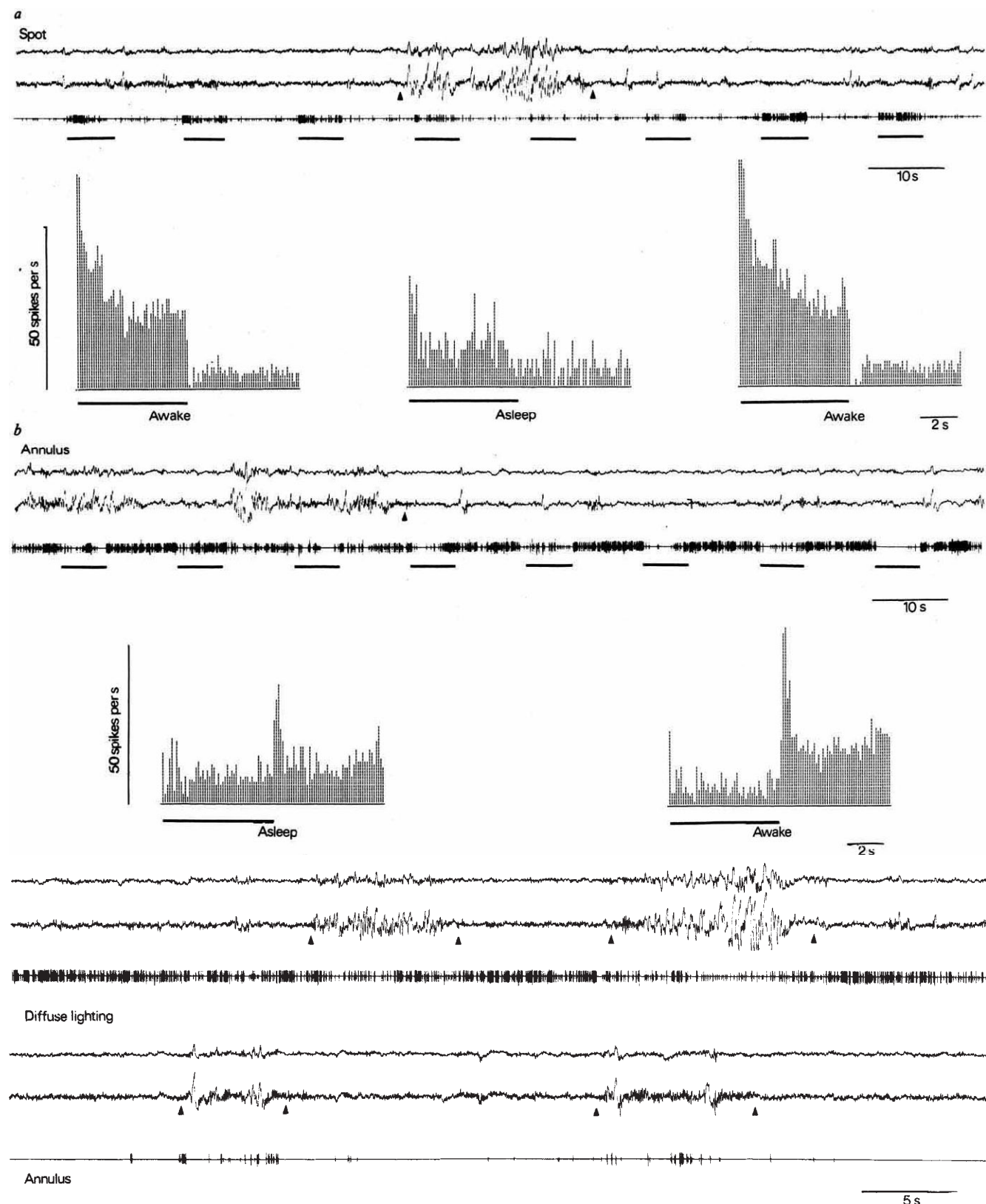


Fig. 1 *a* and *b*, Responses of an on-centre cell from layer A of the lateral geniculate nucleus in a drowsy cat. In each polygraph record the upper two traces are the bipolar EEGs from anterior and posterior head regions; the third trace shows the spikes from the geniculate cell, converted into pulses, in these and subsequent polygraph records, by a Schmitt trigger circuit. Each record lasts ~2 min. Brief periods of sleep, which end spontaneously, are indicated by the slow waves in the middle of record *a* and (roughly) the left half of *b*. The approximate transitions between the flat waking EEG record and periods of slow-wave sleep are indicated by arrowheads. In *a* a spot covering the receptive field centre is flashed on and off. The relative weakness of the on responses during periods of slow waves can be seen in the record of spikes and in the three histograms. The left histogram is the average of 13 responses, of which the last 3 are shown. The middle histogram is an average of the 2 responses during the burst of slow waves, and the right histogram the average of the next 16 responses. In *b* the stimulus is an annulus covering the receptive-field surround. Both the suppression of firing during the stimulus and the off discharge are weaker during sleep. Histograms are the average of nine responses each, before and after the arrowhead. *c*, Effects of slow-wave sleep on steady-state firing of the same cell. In the upper set of three traces (50 s) the screen is diffusely lit by the background light. At each burst of slow waves in the cortical EEG (upper two traces) the rate of firing declines. In the lower set an annulus covers the receptive-field surround. In the aroused periods the discharges are almost completely suppressed, but each episode of slow waves is accompanied by bursts of impulses. Onset and end of slow-wave bursts are indicated by arrowheads.

pressed the firing. The suppression was almost complete when the EEG was flat, but during the slow waves there were bursts of impulses. Thus arousal could depress or elevate the cell's firing rate, depending on the visual stimulus used.

For off-centre cells there was a similar enhancement, both in the suppression of firing and off-discharge to a spot covering the receptive field centre, and in the on-response to an annulus. In both on-centre and off-centre cells diffuse light was less effective than a spot confined to the receptive field centre. This is characteristic of geniculate cells. The diffuse light responses of both on-centre and off-centre cells showed little or no enhancement on arousal from slow-wave sleep, presumably because the increase in the response from the field centre was counterbalanced by an increased opponent response from the surround. To understand the significance of this one must recall that a weaker response to a large spot than to a small one just covering the field centre is characteristic also of retinal ganglion cells^{15,16}. But geniculate cells receiving their main excitatory optic input from a single retinal ganglion cell show a more pronounced reduction in response to stimuli including the field surround than do the retinal ganglion cells^{14,17}; indeed, this is the main known contribution of the lateral geniculate to the processing of visual information. If the response enhancement on arousal were greater for a large spot than for a small one, one would have to conclude that arousal led to an impairment of this function.

Arousal from slow-wave sleep produced no obvious changes in geniculate receptive field sizes as mapped by small spots. As visual acuity is presumably related in geniculate cells to receptive field centre size, an increase in centre size on arousal¹⁸ would have been perplexing if not dismaying.

In summary, arousal from slow-wave sleep increased the spontaneous firing rate of lateral geniculate cells and reduced or

eliminated the high-frequency bursts (not resolved in the slow time scale of Fig. 1); both the excitatory and inhibitory components of responses to spatially restricted visual stimuli were enhanced; and, finally, responses from the receptive field centre and the antagonistic surround were both enhanced.

Presumably the mechanism of these effects depends on non-retinal projections to the LGB because in mammals there do not seem to be projections from brain to retina, and firing of retinal ganglion cells is unaffected by arousal¹. The increase in the on- or off-discharges on arousal is known to reflect an increase in the proportion of optic nerve impulses that succeed in triggering geniculate impulses (the firing index)⁷. The enhanced suppression of discharge seen on arousal may reflect enhanced inhibition from other geniculate neurones that are excited by the stimulus and whose firing index is increased by arousal; or arousal may lead to an enhancement of both excitatory and inhibitory synapses. An increased firing during sleep might also be somehow related to the high-voltage cortical slow waves, which, if present in thalamus and cortex concurrently, could be associated with swings in membrane potential capable of triggering bursts of impulses. Suppressing these slow waves, on arousal, could then abolish the discharges.

Visual cortex

In contrast to geniculate cells, cortical cells varied greatly from one to the next in the degree to which they were influenced by changes in arousal level. Marked differences were often seen between two cells simultaneously recorded from two electrodes spaced 1 mm or less apart, or between two cells simultaneously recorded from one electrode. Nevertheless, each individual cell showed a consistent type of change over several cycles of

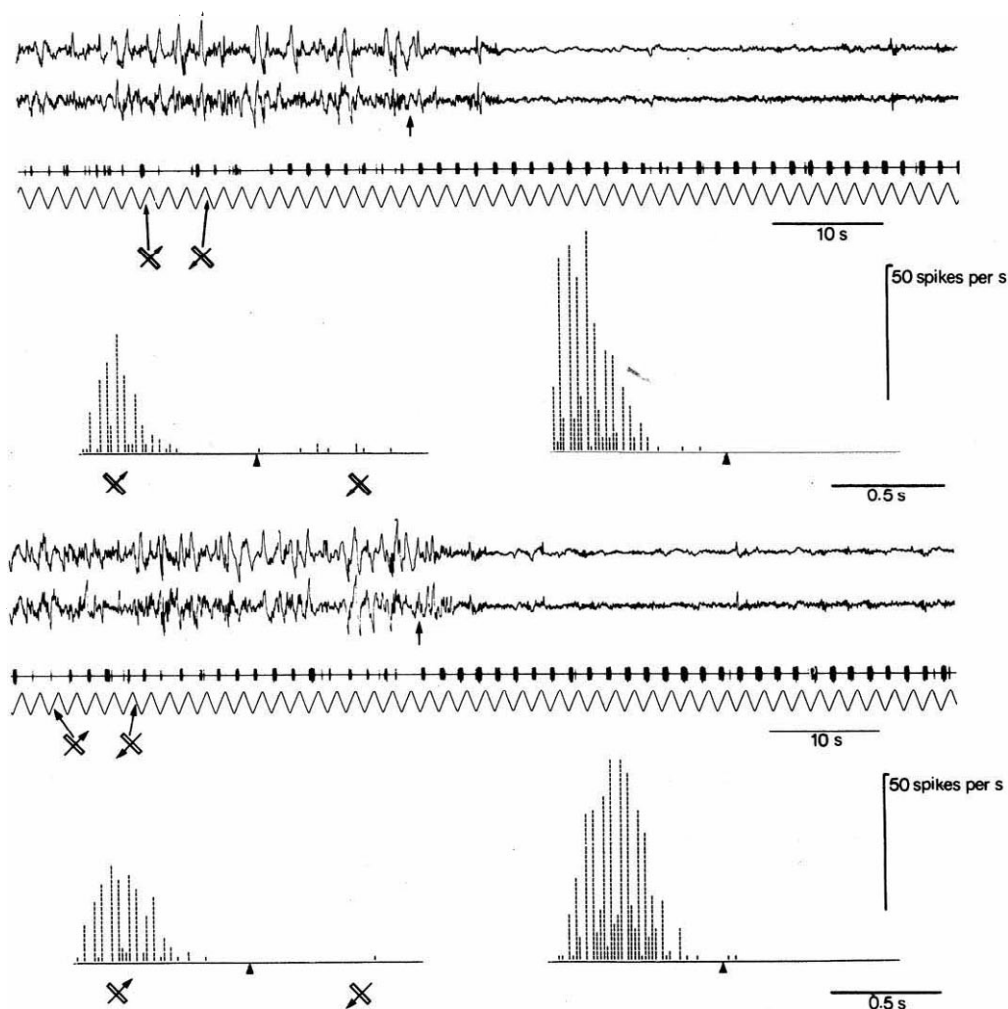


Fig. 2 Effects of arousal on the responses of a cell in layer 6 of striate cortex. The record is 85 s long. Near the middle a noise (arrow) arouses the cat, as shown by the suppression of the EEG slow waves in the upper pairs of traces. Responses of the cell (third trace) are to a $1/2^\circ \times 5^\circ$ stimulus oriented 45° anticlockwise to vertical and swept up and to the right and down to the left (indicated in the fourth trace by up-and-down deflections). On arousal the responses are more consistent and on the whole stronger, as can be seen in the average response histograms. Each histogram is the average of 40 responses. The arrowhead below each histogram indicates when the stimulus reversed direction.

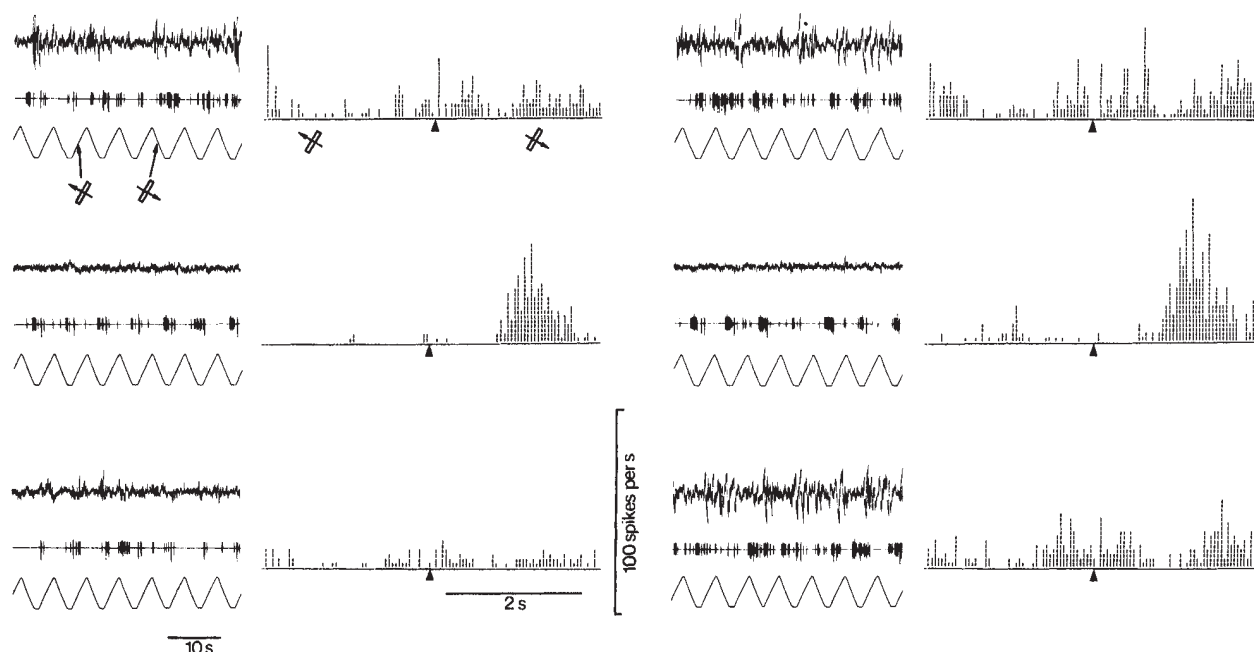


Fig. 3 Responses of a cell in the striate cortex with the cat in various levels of arousal. Each record lasts 45 s and ~3 min elapse between records. As indicated by the EEG (upper trace), the three records on the left side are taken, from above downwards, in slow-wave sleep, awake and slow-wave sleep; those on the right are in slow-wave sleep, awake and slow-wave sleep. The middle trace shows impulse discharges, and lower trace indicates by up-and-down deflections left and right movement of a $1/4^\circ \times 2^\circ$ slit oriented 30° clockwise to vertical. In the alert state the cell responds briskly to rightward movement, but hardly at all to leftward movement. In drowsiness and sleep spontaneous firing is greatly increased, and responses to the moving slit are almost completely absent. Each histogram is the average of 12 responses.

sleeping and waking. Usually (in two-thirds of the 130 cells studied) an irregular burst of spontaneous firing in slow-wave sleep was reduced in overall rate and replaced by a smoother, more regular pattern on awakening. In 10 cells there was an increased firing rate and a smoothing. The rest of the cells showed no difference in spontaneous activity between sleeping and waking. Visually evoked responses were either unaffected or enhanced by awakening: in no cells were responses diminished. Roughly one-third of all cells showed an increase in evoked response that was obvious on the polygraph record; many more cells showed an enhancement of responses that were detected only by histogram averaging. Even in cells showing no increase in evoked response, the response was usually easier to observe against the more regular or reduced background firing.

Figure 2 shows an example of the effects of arousal on the responses of a cortical cell. The stimulus was an optimally oriented moving slit of light. The responses, which were weak and capricious during slow-wave sleep, became stronger and more consistent after arousal by a noise. Despite the increase in response to the optimum direction of movement (up and to the right), the reverse direction remained ineffective.

Figure 3 shows six extracts from a record of another cell, taken over a 15-min period during which the cat repeatedly slept and awoke. The visually evoked responses were drastically changed, from vigorous while awake to almost completely absent during sleep. The increase in the response in the awake state was especially obvious because the background firing was reduced.

Four of the cells showed a transient burst of smooth steady firing lasting for about 30 s when the animal awoke spontaneously or was aroused by a noise or touch. After the burst the firing remained regular, usually stabilizing at a rate about the same as that during sleep or slightly higher. Had these cells not been clearly and specifically driven by particular visual stimuli, they might have been classed as multimodal, given their tendency to respond to any stimulus capable of arousing the animal. We saw no cells in striate cortex that responded to sense modalities other than visual, except in association with arousal.

A fall in both orientation specificity and movement-direction

selectivity has been reported following stimulation of the menencephalic reticular formation¹⁹. Our preliminary results with natural arousal do not, however, point to a loss of specificity; on the contrary there may sometimes be an improvement. Figure 4 shows that the response to an optimally oriented slit was clearly, though not dramatically, improved after arousal, caused in this case by a loud noise, while the background firing was reduced and smoother. Directional selectivity was slightly improved, in that the response to the less effective direction of movement declined. For most directionally selective cells, however, the responses to the two opposite directions changed roughly in proportion. If, as in Fig. 2, one direction of movement failed to evoke a response in sleep, the same was so after arousal. No cell lost directional selectivity after arousal from sleep.

In testing orientation selectivity with a hand-held visual stimulator, we usually found no change. A few cells showed a slight improvement, but to be certain of this it will be necessary to compare orientation tuning curves.

Finally, in several cells we studied the effects of awakening on end stopping—the decline in response that occurs in some cells, termed hypercomplex, when a slit, edge, or dark bar is made longer than optimal²⁰; the degree of stopping did not vary with level of arousal. The enhanced excitation from the activating part of the receptive field is therefore presumably offset by an enhanced inhibition from the outlying areas.

These observations are consistent with previous ones^{20–22} that deepening barbiturate anaesthesia suppresses responses but does not change their specificity or lead to other qualitative changes. They also confirm the finding of Wurtz²³ that cells in area 17 recorded from awake alert monkeys with chronic recording techniques show responses that are not obviously different from responses obtained under barbiturate anaesthesia.

Most of the cells we studied were complex in type. The few simple cells examined showed the same types of changes as the complex cells. Though we have too few cells to make reliable comparisons of laminar distribution, the most striking changes in responses seemed to be in deeper-layer cells.

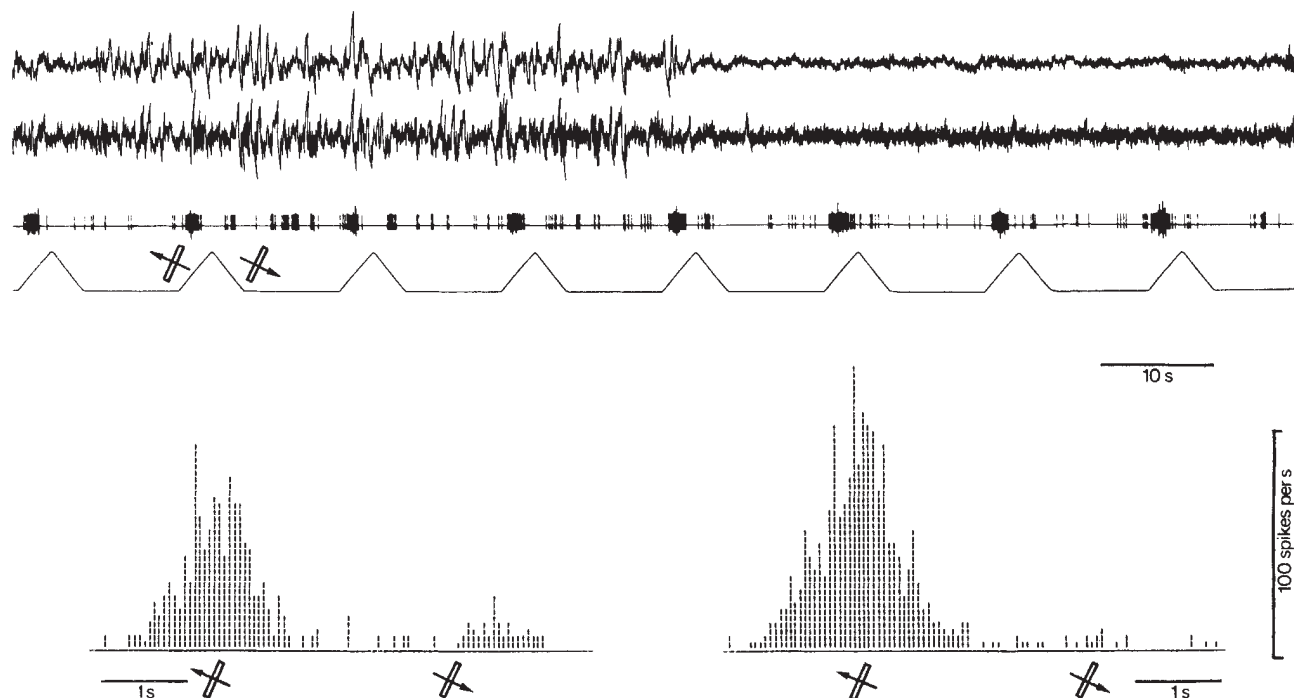


Fig. 4 Effects of arousal from slow-wave sleep on responses and response selectivity of a cell in layer 2 of striate cortex. About halfway through the 2-min record the cat is aroused by a noise. An optimal slit, $1/2^\circ \times 3^\circ$, oriented 25° clockwise to vertical, evokes a response (third trace) that is much greater to movement up and to the left than down and to the right. Arousal results in a moderate increase in the response to leftward movement, and a virtual elimination of the response to rightward movement (see the histograms). Arousal also produces suppression and smoothing of the spontaneous firing.

2-Deoxyglucose studies

We examined deoxyglucose uptake as a function of waking state in eight cats. Four cats, two awake and two in slow-wave sleep, were stimulated with moving vertical stripes confined to the right visual field. The other four cats were half the time awake and half the time asleep (see below). The cats studied in slow-wave sleep were sleep deprived on a treadmill.

Cats were prepared in the same way as those used for physiological recordings, except that the microelectrode advancer was not installed. After discontinuing the anaesthetic we waited until the arrhythmic slow-wave pattern of halothane anaesthesia gave way to the low-voltage rapid activity characteristic of arousal. In the sleep-deprived animals this waking activity was soon replaced by the high-voltage slow waves of slow-wave sleep. A noise or tactile stimulus was sufficient to arouse the animals.

The visual stimulus was a moving set of irregularly spaced vertical black and white stripes projected on a screen by a slide projector and a motor-driven mirror, which produced a back-and-forth movement at 2° s^{-1} . Stripe width was about $1/4^\circ$ near the vertical midline, gradually increasing to $2\text{--}3^\circ$ in the periphery. The stripes were confined to the right field of vision, and did not include the vertical midline but extended from $2\text{--}3^\circ$ to the right of the midline, out to $\sim 45^\circ$. The diffusely lit part of the screen ($0.5 \text{ log cd m}^{-2}$) covered the vertical meridian and left visual field. We injected ^{14}C -2-deoxyglucose in a single pulse over 5–10 s, began the stimulus immediately and continued it for 45 min.

The cats were then deeply anaesthetized with thiopental and perfused with normal saline followed by formol-saline. The brains were removed, frozen in liquid N_2 -cooled Freon-25 at -100°C and sectioned on a cryostat at -30°C . Sections were mounted on glass coverslips, heated to 75°C to drive off water and exposed against X-ray film. We sometimes combined autoradiographs of 2–4 successive cortical sections by making enlargements on film and superimposing them, so that we could average the patterns seen in several single sections and obtain an enhancement of the overall pattern.

Though the two waking brains differed consistently from the two sleeping ones, we were worried that the differences might be due to variability between individual cats rather than to differences in their arousal state. In our first attempt to demonstrate differences in a single animal we alternated the stimuli, presenting stripes to the left visual field when the cat was awake and to the right field when it was asleep. The result was unsatisfactory, perhaps because each hemisphere was stimulated only half the time and any label associated with visual stimuli would be seen against a background representing both states of wakefulness.

To overcome these problems (and also with an eye to countless other applications), we developed a double-label 2-deoxyglucose technique, exploiting the differential sensitivity to ^{14}C and ^3H β -emissions of standard X-ray film and LKB Ultrofilm ^3H . In standard X-ray film the emulsion is protected by a gelatin layer, $\sim 1 \mu\text{m}$ thick, which acts as a partial barrier to the low-energy ^3H emissions. The LKB film lacks this coating and is roughly 20 times more sensitive than X-ray film to ^3H , whereas the two films are about equally sensitive to ^{14}C . These ratios were determined by exposing both types of film to standards made by soaking pieces of filter paper in serial dilutions of ^3H and ^{14}C . In the autoradiographs shown here we increased the ratio of sensitivity to ^{14}C to sensitivity to ^3H of the X-ray film to about 100:1 by interposing an additional 10- μm layer of polyethylene film.

In the injections of deoxyglucose we tried to achieve a $^{14}\text{C}/^3\text{H}$ ratio that would allow the ^3H to overpower the ^{14}C on the LKB film, and the ^{14}C to predominate on the X-ray film.

Both double-label cats were sleep deprived on a treadmill and prepared as described for the single-label (^{14}C) cats. After the halothane was discontinued we waited until the EEG had shifted to a normal slow-wave sleep pattern from which the animal could easily be aroused. We then injected the ^{14}C -2-deoxyglucose and stimulated the left visual field with vertical black and white stripes for 45 min, all the while praying that the animal would stay asleep. Next we aroused the cat, injected the ^3H -deoxyglucose and stimulated the right visual field for 45 min,

keeping the cat awake by a general uproar and tactile stimuli as needed. To be certain of confining the stimulus to the contralateral hemisphere we avoided stimulating a strip of visual field 2–3° to either side of the vertical midline. The cat was perfused and the brain sectioned as already described. The histological sections (together with a set of filter paper standards) were exposed first to X-ray film protected by polyethylene film and then to LKB film.

Lateral geniculate autoradiography

In the two single-label animals whose right visual fields were stimulated when awake, the left LGBs were heavily labelled. The region of dense label was confined to the part of the geniculate corresponding to the area of visual field stimulated, extending on coronal sections medially almost to the medial border, which corresponds to the vertical meridian, and laterally over half to two-thirds of the geniculate's width, corresponding to the 45° lateral extent of the stimulus (Fig. 5a). The right (unstimulated) LGB showed label only at background levels. In one of the two animals stimulated while in slow-wave sleep, both sides showed label only at background levels; in the other animal there was only very faint labelling on the stimulated side (not shown). In the double-label cat (Fig. 5b) only the ^3H -sensitive film showed a left–right difference in geniculate labelling, and again the region with increased label (in the left geniculate) corresponded exactly to the part of the visual field stimulated. No difference was seen between the two sides on the polyethylene-protected X-ray film, except for slightly heightened labelling on the left, in exactly the same place as the label on the ^3H film and doubtless due either to some unfiltered ^3H or to some ^{14}C -deoxyglucose still in the cat's bloodstream 45 min after the first injection, when the second, waking, ^3H stimulus was begun. That the lack of ^{14}C label on the right side was not due to failure of uptake of ^{14}C is clear from the cortical autoradiographs in the same animal (see below). Considering the profound effects of slow-wave sleep on responses of the individual cells we studied, the pronounced differences in the autoradiography are not too surprising.

Cortical autoradiography

A consistent and marked difference was also seen in the pattern of labelling in the visual cortex. Figure 5c shows representative overlaid sections from two single-label (^{14}C) cats, awake and in slow-wave sleep. These animals both had had their right visual fields stimulated, and in each there were conspicuous columns in the left hemisphere, and no hint of any periodic labelling pattern in the right hemisphere. In the awake animal the columns extended through the full thickness of cortex (except for layer 1) with roughly uniform density. In the cats stimulated during slow-wave sleep, the columns in the upper layers (above layer 4) were labelled about as densely as in the waking cats, but below layer 4 they were in most sections hardly visible against the background. In some cats layer 4 was continuously and densely labelled (Fig. 5c, waking) whereas in other cats it was not (Fig. 5c, slow-wave sleep). These differences in layer 4 labelling did not consistently correlate with sleep state and we cannot account for them.

The double-label cat gave the same result (Fig. 6). The upper pair of autoradiograph overlays are made from the same sets of sections as the lower pair, the only differences being in the type of film used and the interposition of a sheet of polyethylene film in the lower (^{14}C) set. Columns extended through the full thickness of cortex in the waking, ^3H -labelled side, and for the most part only halfway down on the sleeping, ^{14}C side. The figure also demonstrates how effectively this technique can distinguish between two different stimuli, as in each case only the appropriate hemisphere shows columns.

Discussion

One might have imagined that when an animal awakens the neurones in its brain would simply become more active, as suggested by Sherrington²⁴. While this may apply to some areas

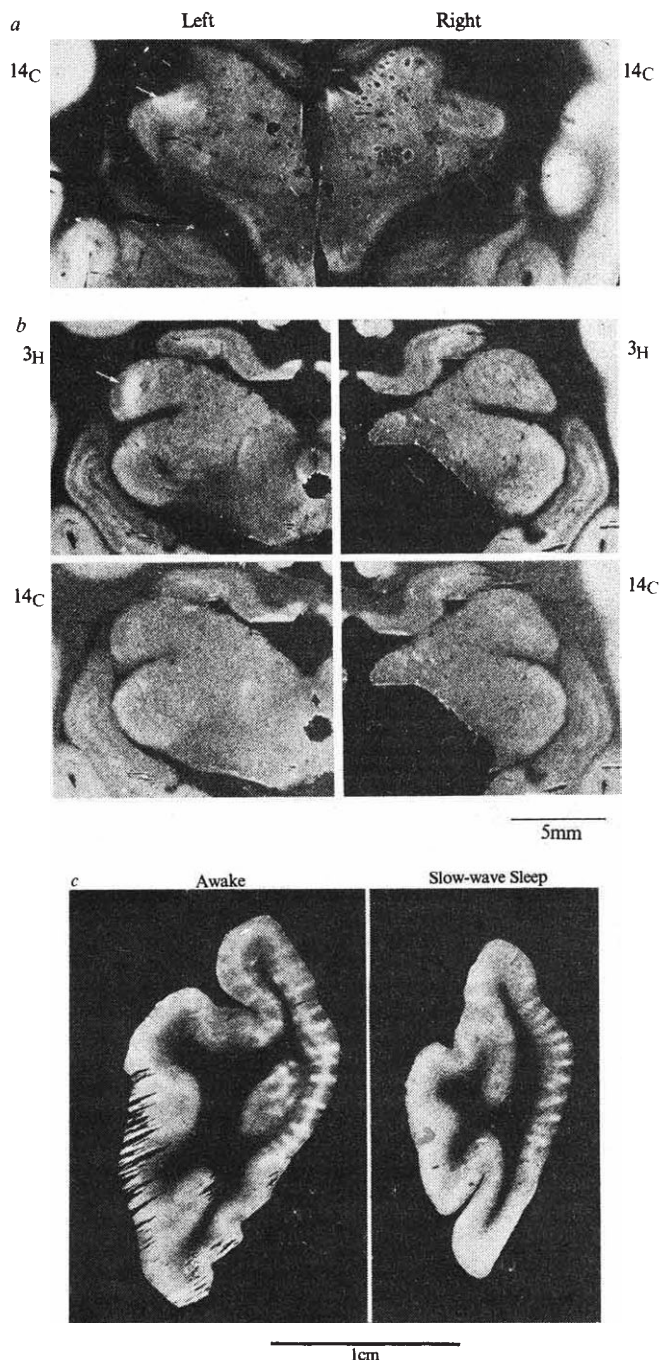


Fig. 5 a, 2-Deoxyglucose labelling of lateral geniculate nucleus in an awake cat by a moving grating of vertical black and white stripes. Stimulus was confined to the right field of vision, coming to within 2° or 3° of the vertical midline. Label in the left LGB is indicated by the arrow. Another waking cat gave a similar result. Figs 5 and 6 are negatives; labelled regions show as bright areas. b, Four autoradiographs showing results of the sleep–waking double-label experiment. After the ^{14}C -2-deoxyglucose injection (100 $\mu\text{Ci kg}^{-1}$) the cat was stimulated while in slow-wave sleep for 45 min, the stimulus being confined to the left visual field. ^3H -2-deoxyglucose (5 mCi kg^{-1}) was then given, and followed by 45 min of stimulation to the right visual field during which the cat was kept aroused. The same sets of sections were exposed to LKB Ultrafilm (^3H) and to polyethylene-protected X-ray film (^{14}C). The ^{14}C autoradiographs of coronal sections show no obvious labelling; the ^3H autoradiographs by contrast show a dense patch of label in the appropriate part of the left LGB. c, Deoxyglucose labelling of visual cortex in two cats, awake and in slow-wave sleep, stimulated with vertical stripes confined to the right visual field. Each part consists of three superimposed successive sections. Columns are labelled on the medial surface (to the right) in both coronal sections through the left hemispheres; the vertical midline projection (postlateral gyrus, at the top) is, as expected, not labelled. In the waking cat the columns (which appear bright in these negatives) extend through the full cortical thickness; in the slow-wave sleep cat they are well labelled in the superficial half of the cortex, but below that they are faint or absent.

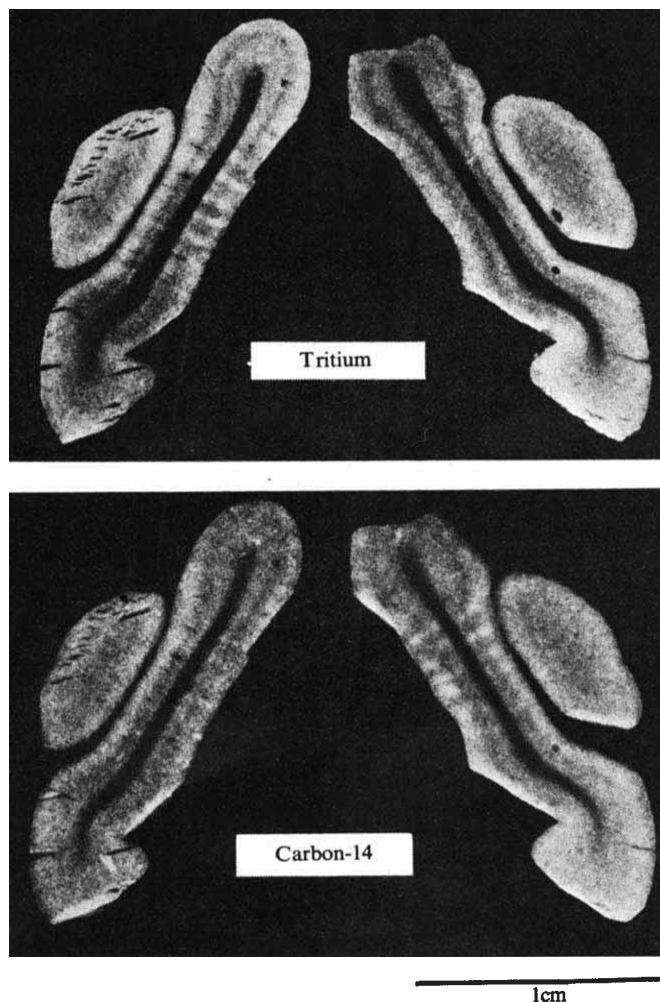


Fig. 6 Double-label deoxyglucose labelling of cortex; three successive sections exposed to LKB film (^3H) and to protected X-ray film (^{14}C) are superimposed. In the upper, waking, ^3H -labelled autoradiograph, columns are seen only in the left hemisphere and extend through the full cortical thickness. In the lower, sleeping, ^{14}C -labelled autoradiograph they are more densely labelled in the upper layers than in the lower ones.

of the brain, for example to parts of the brain stem reticular formation, it does not fit well with what we know of the cerebral cortex. The EEG does not elucidate this question, because although it is a sensitive indicator of arousal, its interpretation in terms of underlying neuronal activity is equivocal: a flattening on arousal from slow-wave sleep can be due to a reduction of activity or a desynchronization; the waves themselves may represent summation of impulses or of graded potentials. Recordings from single cortical neurones have shown that the flattening of the EEG on arousal from slow-wave sleep is most often associated with a lowered rate of spontaneous neuronal firing^{4,5,25}. Some cells, in our experience a minority, show a transient increase in firing rate. Usually high-frequency bursts are replaced by more regular firing.

Early studies of the effects of arousal on responses from the primary visual cortex used direct electrical stimulation at various points along the pathway, or diffuse light flashes. These stimuli have the disadvantage (compared with stimulation by small spots or lines) that they may evoke competing excitatory and inhibitory mechanisms with consequent loss of effectiveness. Nevertheless the responses they do evoke are facilitated by arousal from slow-wave sleep²⁶⁻²⁸. In studies much farther along in the visual pathway, in the inferotemporal region of nitrous oxide-anaesthetized monkeys, neurones have been observed to respond to complex stimuli more vigorously when the EEG is flat than during periods of high-voltage slow waves²⁹.

Our results suggest that arousal brings about an improvement in signal-to-noise ratio: geniculate cells showed enhanced evoked discharges and most cortical cells showed enhanced responses against a lowered or less chaotic background. (Similar effects were seen in somatosensory cortex of monkeys by Gücer³⁰.) But more than that, all components of the response were enhanced, suppression of firing as well as increases. Arousal may thus be attended by either depression or elevation of firing rate, depending on stimulus conditions. A cell in an awake animal is more the slave of incoming sensory inputs; in sleep it is more independent, less fettered by sensory constraints, but more at the mercy of events related to the still all-too-mysterious slow waves.

In normal life, even with the eyes open and surveying a complex scene, only a small fraction of cells in the visual cortex is likely to be engaged at any given instant because, for example, a light-dark contour can have only one orientation, can move only in one direction and at one velocity. For a few cells the discharge rate will probably be increased by the stimulus, but the remainder will continue to fire at their spontaneous rate, which will on the whole be depressed by arousal. If the visual cortex is representative of the rest of the brain, then in wakefulness most cells are not active but held in quiet readiness.

The effects of arousal in the small sample of geniculate cells studied were surprisingly strong, in contrast with the effects on cortical cells, most of which showed only a modest increase in response. This suggests that for some cortical cells the enhanced responsiveness of geniculate cells may be transmitted not as more vigorous firing but as more selective firing. Several cells showed a relatively greater enhancement to an optimal stimulus than to a suboptimal one, and consequently an improved selectivity for orientation or movement direction. We never saw a decline in selectivity.

In the cat the transition from the awake state to slow-wave sleep may be gradual, characterized in the EEG by bursts of high-voltage slow waves similar to those seen in continuous full-blown slow-wave sleep, alternating every few seconds with the low-voltage fast activity of arousal. The EEG alone cannot tell us whether this represents repeated transitory changes in arousal level or instead is a steady-state level of low arousal (drowsiness) characterized by a fluctuating EEG. Our single-cell recordings from geniculate and cortex support the former alternative, given the often tight correlation, within a few seconds, between a cell's behaviour and the onset or termination of an episode of slow waves.

Certain brain-stem neurones have long been known to show firing patterns that are closely correlated with levels of arousal³¹⁻³⁹. Recently attention has focused increasingly on two monoaminergic brain-stem regions, the raphe nuclei and the locus coeruleus. Neurones in both these structures fire more rapidly during periods of increased alertness⁴⁰⁻⁴³. Like visual neurones, the brainstem neurones show a remarkably close correlation with the EEG, but in the brain stem, changes in firing rate usually precede the EEG changes by several seconds^{43,44}. As both pharmacological and physiological studies have suggested that monoaminergic brain-stem neurones are important in controlling arousal, we compared, in a few cats, the effects on visual neurones of stimulating the locus coeruleus to the effects of spontaneous arousal and awakening induced by sensory stimuli. In each case they were the same. The effects of natural arousal from slow-wave sleep on spontaneous and evoked activity could be further exaggerated in an already awake animal by a loud noise or pinch; here again locus coeruleus stimulation had the same effect as a loud noise or a pinch.

The deoxyglucose autoradiographs showed clearly that the depression of activity in slow-wave sleep is different in different layers, being especially profound in 5 and 6. (This is consistent with an observation by Singer *et al.*¹⁹, that cortical cells that could be antidromically activated from tectum or geniculate were particularly likely to be activated by brain-stem stimulation.) Our single-cell recordings had hinted at laminar differences, but it was certainly not true that in sleep suppression

of responses occurred in all of the deeper-layer cells, or in none of the upper-layer cells. The autoradiographic differences may not be due entirely to cell-body labelling: only electron microscopy will show whether the radioactivity is in cell bodies, axon terminals, dendrites, or all three.

The fact that some cells in layers 5 and 6 shut down in slow-wave sleep is pertinent to the functioning of the superior colliculus and LGB, for they are the respective major targets of these two layers. One guess might be that the source of the arousal effects observed in the LGB is the brain stem, because the locus coeruleus is known to send strong projections to the geniculate, and the dorsal raphe nucleus probably does^{45,46}. Arousal could also act on the geniculate indirectly by reviving some otherwise dormant cells in layer 6 (ref. 19). This could be tested by cortical cooling or injection of local anaesthetics during examination of arousal effects in geniculate cells. It will be especially interesting to compare the responses of cells in the superior colliculus during waking and sleeping, because cooling and ablation of the visual cortex are known to abolish directionally selective responses to moving visual stimuli⁴⁷.

Received 5 January; accepted 27 March 1981.

- Hubel, D. H. *J. Physiol., Lond.* **150**, 91–104 (1960).
- Singer, W. *Physiol. Rev.* **57**, 386–420 (1977).
- Hubel, D. H. *Am. J. Ophthalmol.* **46**, 110–122 (1958).
- Hubel, D. H. *J. Physiol., Lond.* **147**, 226–238 (1959).
- Evarts, E. V. in *The Nature of Sleep* (eds Wolstenholme, G. E. W. & O'Connor, M.) 171–182 (Churchill, London, 1961).
- Evarts, E. V., Bental, E., Bihari, B. & Huttenlocher, P. R. *Science* **135**, 726–728 (1962).
- Coenen, A. U. L. & Vendrik, A. J. H. *Expl Brain Res.* **14**, 227–242 (1972).
- Hubel, D. H. & Livingstone, M. S. *Neurosci. Abstr. 10th a. Meet.*, Cincinnati, No. 113.3 (1980).
- Sokoloff, L. *et al. J. Neurochem.* **28**, 897–916 (1979).
- Livingstone, M. S. & Hubel, D. H. *Neurosci. Abstr. 10th a. Meet.*, Cincinnati, No. 113.4 (1980).
- Singer, W. & Dräger, U. C. *Brain Res.* **41**, 214–220 (1972).
- Maffei, L. & Rizzolatti, C. T. *Arch. ital. Biol.* **103**, 609–622 (1965).
- Sakakura, H. *Jap. J. Physiol.* **18**, 23–42 (1968).
- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **155**, 385–398 (1961).
- Kuffler, S. J. *Neurophysiol.* **16**, 37–68 (1953).
- Barlow, H. B. *J. Physiol., Lond.* **119**, 69–88 (1953).
- Cleland, B. G., Dubin, M. W. & Levick, W. R. *Nature new Biol.* **231**, 191–192 (1971).
- Meulders, M. & Godfraind, J. M. *Expl Brain Res.* **9**, 201–220 (1969).
- Singer, W., Tretter, F. & Cynader, M. *Brain Res.* **102**, 71–90 (1976).
- Hubel, D. H. & Wiesel, T. N. *J. Neurophysiol.* **28**, 229–289 (1965).
- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **160**, 106–154 (1962).
- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **195**, 215–243 (1968).
- Wurtz, R. H. *J. Neurophysiol.* **32**, 727–742 (1969).
- Sherrington, C. S. *Man on His Nature*, 183–184 (Doubleday, New York, 1953).

Our physiological findings imply either that the responses of many visual neurones in the central nervous system are enhanced on arousal, or that in sleep the responses are dulled. As the function of sleep is unknown, it is, of course, not clear whether a muffling of sensory input during sleep serves to insulate the animal from its environment to permit uninterrupted sleep or whether sensory systems also need to rest and the decreased responsiveness we see reflects whatever recuperative process the brain undergoes during sleep.

David Freeman did the computer programming and electronics, the histology was done by Birthe Storai, Suzanne Fenstermaker and Debra Hamburger, and the photography by Marc Peloquin. We thank David Ferster for interesting criticisms. Simon LeVay suggested the technique of overlaying adjacent autoradiographs. Some of the early work on developing the double label was done with Dr Gary Blasdel. The work was supported by NIH grant R01 EY00605, grants from the Rowland Foundation and the Klingenstein Fund and by NIH grant R01 NS 07848 to E. A. Kravitz.

- Murata, K. & Kameda, K. *Arch. ital. Biol.* **101**, 306–331 (1963).
- Akimoto, H., Saito, Y. & Nakamura, Y. *Neurophysiologie und Psychophysik des visuellen Systems* (eds Jung, R. & Kornhuber, H.) 363–374 (Springer, Berlin, 1961).
- Evarts, E. V. *J. Neurophysiol.* **26**, 229–248 (1963).
- Evarts, E. V. in *Colloque International sur les Aspects Anato-mo-fonctionnels de la Physiologie du Sommeil*, 189–212 (Éditions du Centre National de la Recherche Scientifique, Paris, 1965).
- Gross, C. G., Rocha-Miranda, C. E. & Bender, D. B. *J. Neurophysiol.* **35**, 96–111 (1972).
- Güger, G. *Expl Brain Res.* **34**, 287–298 (1979).
- Möller, A., Moruzzi, G. & Naquet, R. *EEG J.* **5**, 571–584 (1953).
- Machne, X., Calma, I. & Magoun, H. W. *J. Neurophysiol.* **18**, 547–558 (1955).
- Strumwasser, F. *Science* **127**, 469–470 (1958).
- Schlag, J. *L'activité Spontanée des Cellules du Système Nerveux Central* (Editions Arscia, Bruxelles, 1959).
- Huttenlocher, P. R. *J. Neurophysiol.* **24**, 451–568 (1961).
- Beniot, O. *J. Physiol., Paris* **56**, 259–262 (1964).
- Kasamatsu, T. *Brain Res.* **14**, 506–509 (1969).
- Kasamatsu, T. *Expl Neurol.* **28**, 450–470 (1970).
- Hobson, J. A., McCarley, R. W., Pivik, R. T. & Freedman, R. *J. Neurophysiol.* **37**, 497–511 (1974).
- Chu, N.-S. & Bloom, F. E. *Science* **179**, 908–910 (1973).
- McGinty, D. J., Harper, R. M. & Fairbanks, M. K. *Serotonin and Behavior* (eds Barchas, S. & Usdin, E.) 267–279 (Academic, New York, 1973).
- McGinty, D. J. & Harper, R. M. *Brain Res.* **101**, 569–575 (1976).
- Foot, S. L., Aston-Jones, G. & Bloom, F. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3033–3037 (1980).
- Trulsson, M. E. & Jacobs, B. L. *Brain Res.* **163**, 135–150 (1979).
- Moore, R. Y. & Bloom, F. E. *A. Rev. Neurosci.* **2**, 113–168 (1979).
- Moore, R. Y., Halaris, A. E. & Jones, B. E. *J. comp. Neurol.* **180**, 417–438 (1978).
- Wickelgren, B. G. & Sterling, P. *J. Neurophysiol.* **32**, 16–32 (1969).

LETTERS TO NATURE

Density power spectrum in the local interstellar medium

J. W. Armstrong*, J. M. Cordes† & B. J. Rickett‡

* Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California 91109, USA

† Department of Astronomy, Cornell University, Space Science Building, Ithaca, New York 14853, USA

‡ Department of Electrical Engineering and Computer Sciences, University of California, San Diego, La Jolla, California 92093, USA

Since the identification of interstellar scintillation, the interstellar medium (ISM) has been known to contain electron density fluctuations having scales $\sim 10^9$ m. The scattering properties of these irregularities have been used in pulsar and continuum source studies and to investigate the small scale turbulence itself^{1–11}. (We use 'turbulence' here to imply a disordered, irregular flow.) If the ISM turbulence extends over a broad enough range of fluctuation scales, it could have an important role in galactic cosmic ray transport and confinement^{12–14}. Here we present observational evidence that electron density irregularities exist over a very wide range of scale sizes. Radio scattering observations associated with the wavenumber range $\sim 10^{-11}$ – 10^{-6} m⁻¹ are consistent with a density spectrum of the form (wavenumber)^{-3.7±0.6}. At lower

spatial wavenumbers (10^{-16} – 10^{-18} m⁻¹) the power spectrum can be estimated by computation of the density fluctuations near the 'outer' scale of the spectrum¹⁵, comparison with density irregularities predicted by theoretical models of the ISM¹⁶, and comparison with observations of the velocity structure function^{13,17,18}. When combined with the high-wavenumber (radio) data, these low-wavenumber spectrum estimates are most simply interpreted in terms of a density spectrum behaving as (wavenumber)^{-3.6±0.2} over a 12 decade scale size range (~ 100 pc to $\sim 10^7$ m). There may be theoretical difficulties^{12,14,19} with such a spectrum, however. Alternatively, the low-wavenumber data could be associated with ISM clouds which are physically distinct from the small scale turbulence.

The power spectrum describes the decomposition of a random field's variance into its various Fourier wavenumber components: it partitions the fluctuation 'power' of the field into eddies, with the characteristic size of an eddy being $\sim 2\pi/\text{wavenumber}$. The random field of interest here is the interstellar electron density. We assume that the density spectrum P_{3N} (=Fourier transform of the three-dimensional density covariance function) depends only on the modulus of the wavenumber (isotropic statistics) and is independent of position within the ISM (homogeneous statistics). The homogeneity assumption is expected to break down at scale sizes comparable with the overall dimensions of the flow⁴. We have interpreted observations that are sensitive to particular scale sizes in terms of the density fluctuations required on those irregularity scales.

Table 1 Estimates of P_{3N}

Method	Wavenumber (m^{-1})	Ref.	Theory equation in ref.	Data ref.	Typical value	C_N^2 ($m^{-6.67}$) (if $\alpha = 11/3$)	$P_{3N}(q)$ (m^{-3}) at $q = 10^{-7} m^{-1}$
Angular broadening of:							
Crab pulsar	3.8×10^{-8} to 9.2×10^{-7}	4	9	22, 23	Visibility = 0.32 at $\lambda = 4.05$ m	1.7×10^{-4}	7.9×10^{21}
Crab pulsar	$\sim 10^{-8}$ to $\sim 2.4 \times 10^{-7}$	21	2.8	21, 27, 28	$\theta = 0.1''$ at $\lambda = 4.05$ m	1.2×10^{-4}	5.6×10^{21}
PSR 0834 + 06		4	19	22, 23	$\theta = 0.06''$ at $\lambda = 2.7$ m	4.4×10^{-4}	2.0×10^{22}
PSR 0950 + 08		4	19	22, 23	$\theta = 0.05''$ at $\lambda = 2.7$ m	2.2×10^{-3}	1.0×10^{23}
PSR 1133 + 16	3.8×10^{-8} to 9.2×10^{-7}	4	19	22, 23	$\theta = 0.06''$ at $\lambda = 2.7$ m	1.4×10^{-3}	6.5×10^{22}
PSR 1919 + 21		4	19	22, 23	$\theta = 0.06''$ at $\lambda = 2.7$ m	7.1×10^{-4}	3.3×10^{22}
IPS survey	$\sim 10^{-8}$ to 2.4×10^{-7}	4	19	25, 26	$\theta = 0.12''$ at $\lambda = 3.5$ m	2.4×10^{-4}	1.1×10^{22}
Visibility phase scintillation	1.2×10^{-8} to 3×10^{-7}	4	9	31	$\leq 2 \text{ rad}^2$ at $\lambda = 0.038$ m	≤ 1.5	$\leq 7 \times 10^{25}$ at $10^{-7} m^{-1}$
ISS decorrelation bandwidth	$\sim 6 \times 10^{-9}$ to $\sim 6 \times 10^{-8}$	4	41	1, 3, 11	~ 100 kHz at 0.75 m	3×10^{-5} to 3×10^{-4}	$\sim 10^{26}$ at $6 \times 10^{-9} m^{-1}$
Pulsar timing noise	$\sim 2 \times 10^{-12}$ to $\sim 6 \times 10^{-13}$	4 29	28 198	38–41	≤ 1 ms at $\lambda = 2$ m	≤ 4	$\leq 3 \times 10^{43}$ at $2 \times 10^{-12} m^{-1}$
Frequency time drift for:							
PSR 1133 + 16	$\sim 6.3 \times 10^{-12}$ to $\sim 6.3 \times 10^{-10}$	36, 37		9, 11	$ df/dt \approx 3,000 \text{ Hz s}^{-1}$ at 340 MHz	8.2×10^{-4} at $s = 10^{10}$ m	4×10^{30} at $6 \times 10^{-10} m^{-1}$
PSR 0329 + 54		36, 37		11	$ df/dt \approx 420 \text{ Hz s}^{-1}$ at 408 MHz	9.1×10^{-5} at $s = 10^{10}$ m	5×10^{29} at $6 \times 10^{-10} m^{-1}$
PSR 1642 – 03		36, 37		11	$ df/dt \approx 50 \text{ Hz s}^{-1}$ at 340 MHz	≤ 4 at $s = 10^{10}$ m	$\leq 2 \times 10^{34}$ at $6 \times 10^{-10} m^{-1}$
Velocity structure function	$\sim 3 \times 10^{-18}$ to $\sim 2 \times 10^{-16}$	See text		17, 18	$1 \text{ km}^2 \text{ s}^{-2}$ at 1 pc	3×10^{-4} to 3×10^{-8}	6×10^{56} at $10^{-17} m^{-1}$
Density outer scale	$\sim 2 \times 10^{-18}$	15, 20 and see text		13, 15 17, 49	$L_0 = 100$ pc $\rho = 0.03 \text{ cm}^{-3}$	6×10^{-7}	$\sim 5 \times 10^{58}$
ISM clouds	$\sim 10^{-17}$	16		N/A	0.17 cm^{-3}	N/A	$\sim 10^{59}$

This gives a power spectrum level at $q \sim 2\pi/(\text{scale size})$ that is consistent with observations. Previous studies^{4,8,15} having been confined to fewer types of observational data were limited to estimating the power spectrum in fewer wavenumber ranges. In many wavenumber intervals, we have modelled $P_{3N}(q) = C_N^2 q^{-\alpha}$ where C_N^2 measures the 'level' of the turbulence and α is a parameter (usually such that $2 < \alpha < 4$). The P_{3N} estimates from various wavenumber ranges can then be combined to see if an overall form for the spectrum can be synthesized.

Very small scale structure ($\sim 10^6$ – 10^7 m) can be inferred from observations of the angular broadening of radio sources and very long baseline interferometry (VLBI) phase scintillation. In both observing methods, the phase structure function^{4,20} has a central role. In angular broadening observations an interferometer having baseline b measures the visibility $\Gamma(b) = \exp(-D(b)/2)$, where $D(b)$ is the phase structure function: $D(b) = \langle |\phi(\mathbf{x}) - \phi(\mathbf{x} + \mathbf{b})|^2 \rangle$. For a spherical wave propagating through a turbulent medium located effectively halfway between the source and the observer, $D(b)$ is related to P_{3N} by

$$D(b) = 8\pi^2 z \lambda^2 r_e^2 \int_0^\infty [1 - J_0(qb/2)] P_{3N}(q) q dq \quad (1)$$

where λ is wavelength, z the source–observer distance, J_0 the Bessel function and r_e the classical electron radius. Thus measurement of $\Gamma(b)$ gives in general only an integral constraint on P_{3N} . However, for a fairly wide class of spectra (for example power laws with $2 < \alpha < 4$) the low wavenumber components of P_{3N} do not contribute to the observation. In these cases, the wavenumber range that mainly determines $\Gamma(b)$ depends on P_{3N} : for $P_{3N} \propto q^{-11/3}$ half of the contribution to $D(b)$ comes from the range 0.1 – $2.4 b^{-1}$ while for $P_{3N} \propto q^{-3}$ this range narrows to 1.6 – $5.4 b^{-1}$. IPS (interplanetary scintillation) observations²¹ of the Crab pulsar respond to several effective baselines near 10^6 m and are consistent with $P_{3N} \propto q^{-11/3}$. Hence we associate the spherical wave angular broadening data with the range 0.1 – $2.4 b^{-1}$.

Observations^{22,23} at 111 MHz with $b \sim 2.6 \times 10^6$ m give scattering diameters of 0.05 – 0.07 arc s, yielding estimate of P_{3N} in the range 4×10^{-8} – $9 \times 10^{-7} m^{-1}$ as indicated in Table 1. The IPS method gives information on similar effective baselines²⁴. An

IPS survey²⁵ has been interpreted in terms of interstellar scattering²⁶ and can be incorporated into a power law model framework (see ref. 4 and Table 1). The most detailed IPS observations are of the Crab pulsar^{21,27,28}, as with the IPS survey these data give P_{3N} estimates in good agreement with those derived from VLBI data.

Besides angular broadening, the ISM irregularities can cause fluctuations in the apparent position of a radio source^{4,29,30} which can be measured with an interferometer. Shapiro *et al.*³¹ measured the angular displacement of two radio sources using VLBI, with the observation time short enough that the interferometer was in the same interstellar scintillation (ISS) phase patch³². Repeated trials over ~ 3 yr showed a scatter in the angular distance between the sources of $\sim 2 \times 10^{-9}$ rad. This scatter is dominated by other effects (such as tropospheric scintillation), but can be used to construct an upper limit on the ISM contribution to the interferometric phase fluctuation and hence a limit on P_{3N} in the range 0.05 – $1.2 b^{-1}$ (plane wave). Interpreting 2×10^{-9} rad as an r.m.s., the phase structure function on the $\sim 4 \times 10^6$ m baseline is $\sim 2 \text{ rad}^2$. Using the theory⁴ with a propagation distance^{33,34} ~ 760 pc gives $C_N^2 \leq 1.5 \text{ m}^{-6.67}$ (if $\alpha = 11/3$) and a corresponding P_{3N} limit that is consistent with the above angular broadening measurements.

In principle, pulse broadening^{2,6} can also be used to estimate P_{3N} at these large wavenumbers. However, several kpc long propagation distance is required to obtain a measurable effect. Because the ISM turbulence is inhomogeneous on these scales we do not try to incorporate pulse broadening data here.

For scales $\sim 10^9$ m the simplest way to sense ISM eddies is through intensity fluctuations in time^{5,35} on frequency-time^{1,3,9,11}. In this method, ISM eddies larger than about the Fresnel zone size do not contribute to the observation^{4,21,24,29}; signal-to-noise considerations effectively limit the extent to which very small scales can be sensed. By measuring the auto-correlation of the intensity fluctuations in radio frequency, one can get the level and shape of P_{3N} near $q = 10^{-8} m^{-1}$. These observations¹¹ favour power law models with α slightly less than 4. Table 1 shows the range of P_{3N} that are consistent with the measured ISS decorrelation bandwidths of relatively nearby pulsars.

Scale sizes between $\sim 10^{10}$ and 10^{12} m can be detected either through time-of-arrival, fluctuations of pulsar pulses^{4,29} or, indirectly, through their effect on the frequency-time structure of the intensity ISS caused by smaller scales^{36,37}. The time of arrival fluctuations are caused by time-varying group delay as the blobs move through the pulsar-Earth line. Careful timing observations³⁸⁻⁴⁰ have yielded only a ~ 1 ms upper limit to ISS time of arrival fluctuations at $\lambda \approx 2$ m and on time scales $\sim 1-3$ yr (that is typical spatial scales $3 \times 10^{12}-10^{13}$ m). For a propagation distance of 1,000 pc we obtain P_{3N} limits as indicated in Table 1.

Alternatively, one can estimate P_{3N} on these wavenumbers through the influence of the large scale blobs on the small scale intensity ISS. The theory is best formulated for two discrete scales³⁶ and we use it to obtain consistent values of P_{3N} near $q = 10^{-11} \text{ m}^{-1}$. The physical idea is that, owing to transverse gradients in the phase, the large blobs cause refractive steering of the intensity diffraction pattern of the small ($\sim 10^9$ m) blobs. The frequency dependence of the plasma refractive index means that there will be a frequency dependent spatial offset of the intensity pattern. If the pattern moves across the Earth with a constant velocity (for the ~ 1 h of a typical observation) these spatial offsets at different frequencies are manifested as temporal offsets in a frequency-time spectrum (see Fig. 2 of ref. 11). Theory allows computation of the angle of refraction, θ_R , of the large blobs from the slope (df/dt) of the drifting features in frequency-time. Because we are dealing with a discrete scale model, we associate θ_R with the transverse phase gradient³⁶ ($\theta_R = (\lambda/2\pi) \cdot |\text{grad } \phi|$) and then to square-law phase structure function²⁹ $|\text{grad } \phi| \approx D^{1/2}(b)/b$, where b is a typical refractive blob scale size. The structure function's relationship to P_{3N} is shown for three pulsars in Table 1. Because our theory is really only for two discrete scales the existence of drifting features does not necessarily imply a continuum of scales between $\sim 10^{10}$ and 10^{12} m.

Some observational aspects of these high-wavenumber P_{3N} estimates should be noted. First, differential VLBI is within two orders of magnitude of being limited by ISS in relative source position measurements. Second, searches for gravitational radiation having periods 1-10 yr using pulsar timing data⁴¹ are apparently limited by pulsar timing noise rather than ISS. (In the Solar System, searches for 30-3,000-s period gravity waves^{42,43} are limited by interplanetary phase scintillation⁴⁴.) Third, in addition to dispersion measure time of arrival fluctuations, the two-scale theory³⁶ admits the possibility of time of arrival variations $\tau_R \approx z\theta_R^2/2c$ (z = typical distance to a scattering element) associated with the changing 'ray' path length to the pulsar. If $\theta_R \sim \theta_s$ (the diffraction scattering angle) then τ_R can be comparable with $z\theta_s^2/2c$ ($\sim$ a few tens of milliseconds for a highly dispersed pulsar) and hence observable.

The radio data alone are consistent with a pure power law spectrum, $P_{3N} \propto q^{-3.7 \pm 0.6}$ for $10^{-11} \leq q \leq 10^{-6} \text{ m}^{-1}$. This includes both the Kolmogorov spectrum ($\alpha = 11/3$) and the $\alpha = 7/2$ spectrum predicted by dimensionality relations for magnetoacoustic turbulence^{45,46}. Both spectra imply a cascade of energy between wavenumbers as in a turbulence process. However, the radio data alone cannot exclude the case $\alpha = 4$ for which the physical explanation is very different. A random assembly of density discontinuities would contribute a q^{-4} spectrum at high wavenumbers; such a spectrum does not imply any energy cascade. However, where sufficiently accurate ISS radio data have been recorded¹¹, the spectra near a wavenumber $q \sim 10^{-9} \text{ m}^{-1}$ are consistent with a power slightly less than 4. Further, if the estimates near $q \sim 10^{-18} \text{ m}^{-1}$ discussed below are correct, the data are consistent with a (wavenumber)^{-3.6 $\pm$ 0.2} power law over 12 orders of magnitude in wavenumber and exclude the $\alpha = 4$ case.

Low wavenumber estimates can be obtained by computing the density fluctuations near the 'outer' scale of the turbulence, by interpreting the velocity structure function data, and through theoretical predictions for the large-scale density fluctuations in the ISM. The outer scale argument is quite general^{15,20} and as the density structure function $D_N(r)$ is the mean square of the

difference in density at two points, $D_N(L_0)$ is expected to be comparable to the square of the difference in average density at points separated by the outer scale size L_0 , thus $D_N(L_0) \sim (L_0 \text{ grad } \bar{\rho})^2$. If $D_N(r) = (C_N^2/0.033)r^{2/3}$ then $C_N^2 \sim 0.033 L_0^{4/3} (\text{grad } \bar{\rho})^2$, where the factor 0.033 is a consequence of our spectrum definition (see refs 4 and 20). Taking $\bar{\rho} = 0.03 \text{ cm}^{-3}$, $L_0 = 100 \text{ pc}$ (refs 15, 17), and the gradient scale for $\bar{\rho}$ to be 500 pc (ref. 34) gives $C_N^2 \sim 10^{-6.2} \text{ m}^{-6.67}$. The associated value of $P_{3N}(2\pi/L_0)$ is plotted in Fig. 1 with an error of $\pm$ two orders of magnitude.

Velocity structure function data give further evidence for a turbulence-type power law spectrum at low wavenumbers. The velocity structure function $D_v(r) = \langle [v(s) - v(s+r)]^2 \rangle$ is measured^{17,18} from the motion of both stars and neutral gas to be $D_v \propto r^{2/3}$ (which corresponds to the Kolmogorov spectrum) for scales 1-100 pc (Fig. 16 of ref. 17 and Fig. 1 of ref. 18 are particularly striking in this regard). To relate the level of the velocity fluctuations to the level of density fluctuations (hence P_{3N}) we have to consider a model of the ISM. There is now strong evidence for a coronal phase (10^6 K, diffuse, ionized) in the ISM⁴⁷. Our interpretation is based on a specific model¹⁶ with three main components: cool ($\sim 10^2$ K) dense neutral gas clouds surrounded by a warm ($\sim 10^4$ K) gas partially ionized by stellar UV, embedded in the hot coronal phase driven by supernova expansions. The coronal phase occupies a large fraction of the space but only contributes a small fraction of the ionized column density on typical pulsar distances. Thus at low wavenumbers $P_{3N}(q)$ probably reflects the warm phase which is thought to be closely associated with the cool neutral phase. The neutral gas motions, $D_v \propto r^{2/3}$, are also typical of the warm ionized phase and imply a Kolmogorov spectrum shape near $q \sim 3 \times 10^{-18} - 2 \times 10^{-16} \text{ m}^{-1}$. It is not clear, however, how to relate the magni-

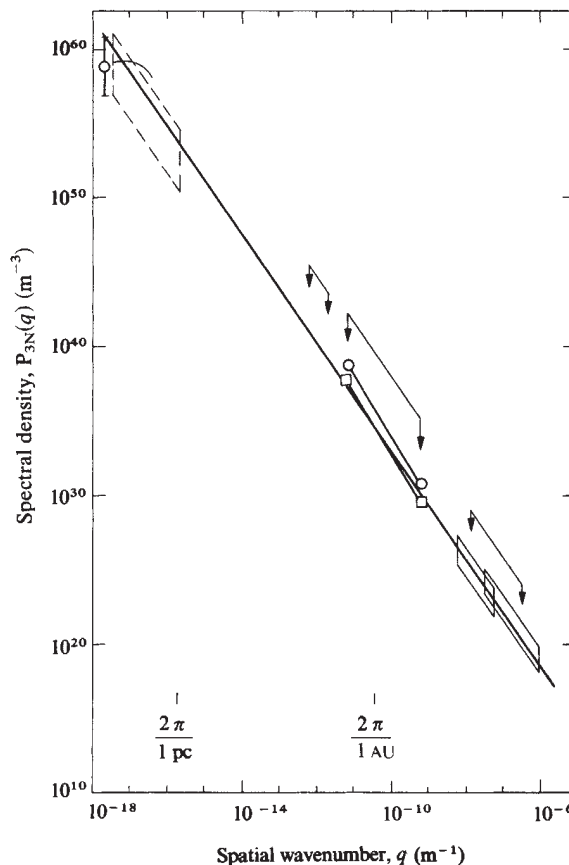


Fig. 1 Density power spectrum in the local interstellar medium. Spectral estimates are obtained from Table 1. Starting at the lowest wavenumber, estimates or limits come from: density outer scale, ISM clouds, velocity structure function, pulsar timing noise, frequency-time drift in dynamic spectra, intensity scintillation, visibility phase scintillation, and angular broadening. The line through the data has a logarithmic slope of $-11/3$ (Kolmogorov spectrum); slope of -3.5 is also consistent with the data.

tude of the velocity fluctuations to the level of P_{3N} for the density fluctuations in the warm partially ionized plasma. The estimate included here is based on analogy with the solar wind plasma, but it may not be physically justified for the ISM. For scales which are small compared with the overall dimensions of the solar wind flow, the fractional density fluctuations are proportional⁴⁸ to the velocity fluctuations divided by the Alfvén speed: $|\delta\rho/\bar{\rho}| = a|\delta v/v_A|$ where $a = 10^{0\pm1}$. We have converted these data to P_{3N} as shown in Table 1, using $\bar{\rho} = 0.03 \text{ cm}^{-3}$ and $v_A = 10 \text{ km s}^{-1}$. Because of the uncertainty in relating the levels of the velocity and density fluctuations, the P_{3N} estimates are shown in Fig. 1 as a dashed box. The vertical extent of the box is the $\pm$ two decade uncertainty associated with a .

Another low-wavenumber estimate is based on the density inhomogeneities predicted by the above ISM model. The model predicts that a significant fraction of space is filled with warm partly ionized cloud shells having typical electron densities $\sim 0.17 \text{ cm}^{-3}$ and typical radii $\sim 2.1 \text{ pc}$. If these are approximated as spheres having gaussian radial density profiles, we obtain the spectrum indicated in Fig. 1, plotted over a decade of wavenumber near $q = 10^{-17} \text{ m}^{-1}$. This estimate agrees well with both the outer scale point and the spectrum inferred from the velocity data.

The striking feature of Fig. 1 is that the data are most simply represented as proportional to $q^{-\alpha}$ with $3.4 \leq \alpha \leq 3.8$ over 12 orders of magnitude in q . These conclusions are in general agreement with those of other workers^{13,15,37,49}, although the present data set is more extensive. However, there may be physical difficulties with an inertial subrange-type spectrum over this very wide scale size range. The excitation and damping of various wave modes in the ISM have been discussed elsewhere^{12,14,19,48,51}. It is expected that the coronal phase could¹² support a power law-type spectrum between $10^{-17} \text{ m}^{-1} \leq q \leq 10^{-9} \text{ m}^{-1}$, with higher wavenumber modes damped. The warm partly ionized component, on the other hand, is thought¹⁴ only capable of supporting fluctuations with $q \leq 10^{-13} \text{ m}^{-1}$. Finally, the cool nearly neutral component is predicted¹⁹ to contain only high-wavenumber ($q \geq 10^{-9}$) irregularities. If these theoretical difficulties are real, different wavenumber ranges in Fig. 1 must be associated with different ISM components. High-wavenumber pulsar scintillations and the pulsar dispersion measure might then be linked only statistically through the distance. An alternative picture⁵⁰ of the ISM associates dispersing electrons with medium density Strömgren spheres of a space density much lower than that of the warm clouds in the model¹⁶ discussed above. Further investigations into the scale heights of the dispersing and (high-wavenumber) scattering electrons will reveal whether they exist in the same phase. If different ISM components contribute to different wavenumber ranges of P_{3N} , the very low wavenumber Fourier components might be associated with warm, partially ionized clouds that truly have a characteristic scale. The large scale fluctuations could then be physically distinct from the high wavenumber turbulence. This view requires an approximately power law spectrum over a much smaller range of wavenumbers and thus partially obviates the theoretical objections noted above. The lack of observations in the 10^{-16} – 10^{-12} m^{-1} range does not rule out this interpretation; however, it seem at odds with the continuum of scales suggested by the shape of the velocity structure function.

Our estimate of a global average wavenumber spectrum for the interstellar electron density is most simply represented by a power law $\propto q^{-\alpha}$ with $3.4 \leq \alpha \leq 3.8$ over $10^{-18} \leq q \leq 10^{-6} \text{ m}^{-1}$, suggesting energy exchange over 12 orders of magnitude in scale. There are theoretical arguments against such a view based on current studies of wave modes in the warm and hot phases of the ISM. Clearly future work must reconcile the results of Fig. 1 with models of the ISM and also examine the possibility of energy exchange between irregularity scales over a very wide range.

We thank J. R. Jokipii for interesting discussions and for bringing ref. 17 to our attention. This work was supported in part by the Caltech President's Fund, NASA contract NAS

7-100, NSF grant ATM 8012801, and the National Astronomy and Ionosphere Center which is operated by Cornell University under contract with the NSF.

Received 4 March; accepted 16 April 1981.

1. Rickett, B. J. *Mon. Not. R. astr. Soc.* **150**, 67–91 (1970).
2. Sutton, J. M. *Mon. Not. R. astr. Soc.* **155**, 51–64 (1971).
3. Roberts, J. A. Ables, J. G. Preprint (1980).
4. Rickett, B. J. *A. Rev. astr. Astrophys.* **15**, 479–504 (1977).
5. Backer, D. C. *Astrophys. J.* **190**, 667–671 (1974).
6. Komesaroff, M. M., Hamilton, P. A. & Ables, J. G. *Aust. J. Phys.* **25**, 759–777 (1972).
7. Rickett, B. J. & Lang, K. R. *Astrophys. J.* **185**, 945–950 (1973).
8. Mutel, R. L. *et al. Astrophys. J.* **193**, 279–282 (1974).
9. Ewing, M. S., Batchelor, R. A., Friefeld, R. D., Price, R. M. & Staelin, D. H. *Astrophys. J.* **162**, L169–L172 (1970).
10. Slee, O. B. *et al. Mon. Not. R. astr. Soc.* **167**, 31–47 (1974).
11. Armstrong, J. W. & Rickett, B. J. *Mon. Not. R. astr. Soc.* **194**, 623–638 (1981).
12. Cesarsky, C. J. *A. Rev. Astr. Astrophys.* **18**, 289–319 (1980).
13. Jokipii, J. R. *Interaction of Charged Particles with Random Electromagnetic Fields in Astrophysics* (Summer school on Plasma Instabilities and Turbulence, Aix-en-Provence, 1974).
14. McIvor, I. *Mon. Not. R. astr. Soc.* **179**, (1977); **178**, 85–99 (1977).
15. Lee, L. C. & Jokipii, J. R. *Astrophys. J.* **206**, 735–743 (1976).
16. McKee, C. F. & Ostriker, J. P. *Astrophys. J.* **218**, 148–169 (1977).
17. Kaplan, S. A. *Interstellar Gas Dynamics* (Pergamon, Oxford, 1966).
18. Larson, R. B. *Mon. Not. R. astr. Soc.* **186**, 479–490 (1979).
19. Scheuer, P. A. G. & Tsytovich, V. N. *Astrophys. Lett.* **7**, 125–127 (1970).
20. Ishimaru, I. *Wave Propagation and Scattering in Random Media* Vol. 2, 538 (Academic, New York, 1978).
21. Kaufman, J. J. thesis, Univ. California (1976).
22. Vandenberg, N. R. thesis, Univ. Maryland (1974).
23. Resch, G. M. thesis, Florida State Univ. (1974).
24. Cohen, M. H., Gundermann, E. J., Hardebeck, H. E. & Sharp, L. E. *Astrophys. J.* **147**, 449 (1967).
25. Readhead, A. C. S. & Hewish, A. *Mem. R. astr. Soc.* **78**, 1 (1974).
26. Duffett-Smith, P. J. & Readhead, A. C. S. *Mon. Not. R. astr. Soc.* **174**, 7–17 (1976).
27. Coles, W. A. & Kaufman, J. J. *Mon. Not. R. astr. Soc.* **181**, 57P–59P (1977).
28. Bell, S. J. & Hewish, A. *Nature* **213**, 1214 (1967).
29. Lovelace, R. V. E. thesis, Cornell Univ. (1970).
30. Cronyn, W. M. *Astrophys. J.* **174**, 181–200 (1972).
31. Shapiro, I. I. *et al. Astr. J.* **84**, 1459–1469 (1979).
32. Cohen, M. H. & Cronyn, W. M. *Astrophys. J.* **192**, 193–197 (1974).
33. Taylor, J. H. & Manchester, R. N. *Astrophys. J.* **215**, 885–896 (1977).
34. Readhead, A. C. S. & Duffett-Smith, P. J. *Astr. Astrophys.* **42**, 151–153 (1975).
35. Downs, G. S. & Reichley, P. E. *Astrophys. J. Lett.* **163**, L11–L16 (1971).
36. Shishov, V. I. *Soviet Astr.* **17**, 598–602 (1974).
37. Hewish, A. *Mon. Not. R. astr. Soc.* **192**, 799 (1980).
38. Helfand, D. J., Taylor, J. H., Backus, P. R. & Cordes, J. M. *Astrophys. J.* **237**, 206–215 (1980).
39. Cordes, J. M. & Helfand, D. J. *Astrophys. J.* **239**, 640–650 (1980).
40. Gullahorn, G. E. & Rankin, J. M. *Astr. J.* **1219** (1978).
41. Detweiler, S. *Astrophys. J.* **234**, 1100–1104 (1979).
42. Estabrook, F. B. & Wahlquist, H. D. *Gen. rel. Gravit.* **6**, 439 (1975).
43. Wahlquist, H. D., Anderson, J. D., Estabrook, F. B. & Thorne, K. S. *Atti dei Convegni Lincei*, **34**, 335 (1977).
44. Armstrong, J. W., Woo, R. & Estabrook, F. B. *Astrophys. J.* **230**, 570–574 (1979).
45. Kraichnan, R. H. *Phys. Fluids* **8**, 1385 (1965).
46. Chashei, I. V. & Shishov, V. I. *Geomagn. Aeron.* **17**, 662–667 (1977).
47. McCray, R. & Snow, T. P. *A. Rev. Astr. Astrophys.* **17**, 213–240 (1979).
48. Neugebauer, M., Wu, C. S. & Huba, J. D. *J. geophys. Res.* **83**, 1027–1034 (1978).
49. Jokipii, J. R. & Lerche, I. *Astrophys. J.* **157**, 1137 (1969).
50. Salpeter, E. E. in *The Large Scale Characteristics of the Galaxy* (ed. Burton, W. B.) 245–252 (Reidel, Dordrecht 1979).
51. Hall, A. N. *Mon. Not. R. astr. Soc.* **190**, 353–369 (1980).

Atmospheric nitrous oxide produced by solar protons and relativistic electrons

Sheo S. Prasad

Jet Propulsion Laboratory, California Institute of Technology,
4800 Oak Grove Drive, Pasadena, California 91109, USA

E. C. Zipt

University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

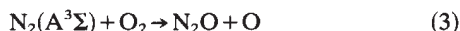
Nitric oxide (NO), which contributes to the destruction of stratospheric ozone, may be formed directly in the upper atmosphere by solar protons^{1,2} and by the precipitation of relativistic electrons from the Earth's radiation belts³. We now describe an alternative means by which solar proton (SP) events and relativistic electron precipitation (REP) events may lead to the production of stratospheric NO—the production of nitrous

oxide (N_2O) in the mesosphere, its downward migration and conversion in the stratosphere to NO by the reaction



This process could amplify the direct NO production by $>10\%$, which is significant. Mesospheric nitrous oxide mixing ratios increase to values as high as 6×10^{-7} due to REP- and SP-related production. Lateral transport will reduce these high values. But even so, mesospheric mixing ratios of N_2O in the high latitudes would approach 10^{-7} , which is considerably greater than those expected on the basis of theories which neglect REP- and SP-related production of this species.

Production of N_2O by precipitating energetic particles is a consequence of the following two reactions⁴⁻⁸.



where e^- represents the primary and secondary electrons. Although the amount of $\text{N}_2(\text{A}^3\Sigma)$ per ion pair is uncertain, a value of 0.52 $\text{N}_2(\text{A}^3\Sigma)$ per ion pair in air seems most probable on the basis of most recent electron impact cross-sectional data of Cartwright *et al.*⁹ as interpreted by Gilmore¹⁰. Zipf⁵ has experimentally determined a 60% probability for N_2O production in the quenching of $\text{N}_2(\text{A}^3\Sigma)$ by O_2 which is in good agreement with earlier experimental results of Malcombe-Lawis⁴. Furthermore, chemical reaction (2) and physical quenching by O_2 seem to be the sole fate of $\text{N}_2(\text{A}^3\Sigma)$ below about 90 km (refs 7, 8). Thus, a value of 0.31 N_2O per ion pair has been adopted in this study.

Production of N_2O by energetic particles can now be obtained directly from the ion production rates due to these particles. Instead of considering an isolated SP or REP event, we are considering the collective effects of all particle precipitation events, both large and small. This has been done in a simplified manner by using yearly averages of ion production rates. Figure 1 shows the results for energetic solar protons and relativistic electrons. Production rates for solar protons are based on the average ion production rate calculated by C. H. Jackman (personal communication) for the year 1978 on the basis of SP observations by experiment on board IMP 8. For relativistic electrons we used the estimates of ion production rates given by Thorne¹¹. We chose to use the ion production rate corresponding to REP occurrence probability of 10% rather than 1%, because recent experiments using heavy-duty magnetic spectrometers with a large geometric factor flown on S3-2 and bremsstrahlung X-ray imaging experiments flown on 1972-076B satellites suggest that the occurrence probability, as well as the fractional area of the globe affected by REP events, are significantly larger than their current estimates¹². Thorne's¹¹ ion production rates are for the subauroral zone ($4 \leq L \leq 10$). Possibly similar production rates may be valid for some higher latitudes also. Nitrous oxide production rates used in our calculations are the sums of the two rates shown in Fig. 1.

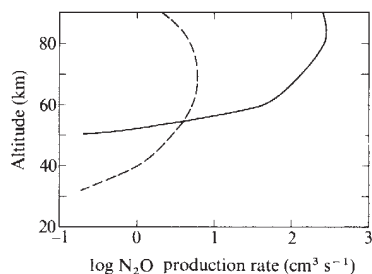


Fig. 1 N_2O production rates: ---, average value for the year 1978 due to solar protons; —, due to relativistic electrons assuming REP events occur 10% of the time.

Atmospheric number densities and mixing ratios of N_2O corresponding to these sources were calculated using one-dimensional photochemical-dynamical modelling techniques¹³. We assumed that N_2O molecules are lost by photodissociation and by reaction with $\text{O}(^1\text{D})$, and by ionic reactions in the ionosphere. Concentrations of $\text{O}(^1\text{D})$ were calculated on the basis of local chemical equilibrium between its production from O_3 photolysis and loss by quenching collisions with N_2 and O_2 . The empirical method of Shimazaki *et al.*¹⁴ was used to determine the absorption of solar UV in the Schuman-Runge bands. We assumed an N_2O mixing ratio of 3×10^{-7} at the lower boundary and a zero flux at the upper boundary at 100 km. Our model was verified by making it reproduce currently believed theoretical distributions¹⁵ of N_2O in the high latitudes ($\sim 60^\circ$) when the flux from the Earth's surface is assumed to be the only source. Results from calculations done for a subauroral latitude of 65° and equinoctial conditions are illustrated in Figs 2 and 3.

From the data presented in Fig. 2, it is clear that nitrous oxide production by REP and SP events could lead to substantial enhancements in high-latitude stratospheric and mesospheric N_2O mixing ratios. These mixing ratios reach values as high as

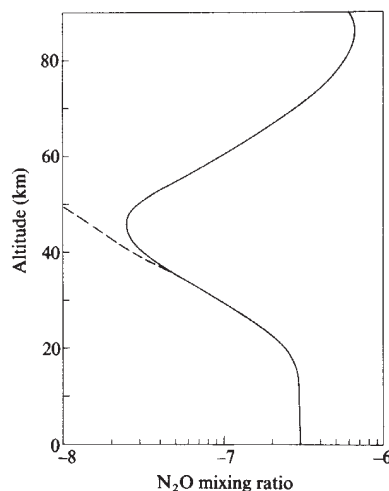


Fig. 2 Theoretical N_2O mixing ratios in the high latitudes neglecting horizontal transport: —, with the N_2O production due to energetic particles; ---, without this source.

$(3-5) \times 10^{-7}$ in the 70–80 km region in the present one-dimensional calculation. Introduction of meridional transport will reduce these numbers. But even so, mesospheric N_2O mixing ratios in the high latitudes would remain substantially greater than those expected on the basis of theories which neglect REP- and SP-related production. This was verified by an approximate calculation in which N_2O loss rate coefficients were systematically increased to simulate roughly the effects of meridional flow.

Indirect column production rates of NO for the various altitudes (h) through reaction (1), the integral $2k_1 \int_h^{100\text{ km}} n(\text{N}_2\text{O})n(\text{O}^1\text{D}) dh$, are shown in Fig. 3, where curves *a* and *b* represent the column production rates for the two cases of with and without the particle precipitation related N_2O source. The indirect production of NO through reaction (1) should be distinguished from the direct production of NO discussed by Crutzen *et al.*¹ and Thorne³. The column production rates for this direct case are shown by curve *c* in Fig. 3 for comparison with the indirect case. In either case, the production rate of NO in any given region of the atmosphere bounded by altitudes h_1 and h_2 ($h_2 > h_1$) can be obtained by simply subtracting the column production rate at h_2 from that at h_1 . In this manner, the data presented by curves *a* and *b* of Fig. 3 show that the indirect

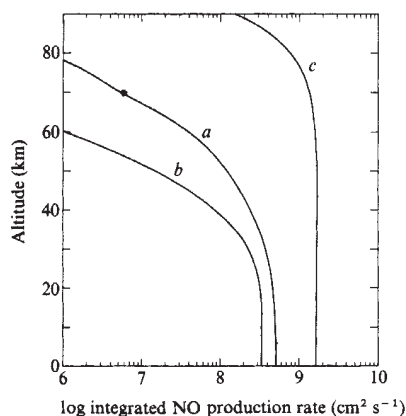


Fig. 3 Column production rate of NO. *a*, Indirect production using N_2O profile obtained with the energetic particle related source of nitrous oxide; *b*, the same without this source; *c*, for direct production of NO by energetic particles.

production of NO in the stratosphere and mesosphere due to precipitating particles was $1.6 \times 10^8 \text{ cm}^{-2} \text{ s}^{-1}$ in our model. Direct production of NO in the same column was $1.7 \times 10^9 \text{ cm}^{-2} \text{ s}^{-1}$ (see curve *c*). Thus, the N_2O mechanism of indirect NO production amplified the direct NO production by $\sim 10\%$. Furthermore, most of the indirect production of NO is below 60 km, that is, $1.2 \times 10^8 \text{ cm}^{-2} \text{ s}^{-1}$ out of the total $1.6 \times 10^8 \text{ cm}^{-2} \text{ s}^{-1}$ or 75%. In contrast, 75% of the direct production is above 72 km in the present model, where photodissociation rates are higher compared with the rates at and below 60 km (ref. 16). Consequently, only a part of the directly produced NO can be transported below 60 km. Thus, the indirect production of NO through the N_2O mechanism could be more significant than what is implied by the 10% amplification mentioned earlier.

Finally, note that present methods of remotely measuring N_2O from the satellite platform, such as used by the Stratospheric and Mesospheric Sounder¹⁷ on board Nimbus-7, are not considered effective above about 55 km, where measurements are needed most to validate the present theory. Alternative means of measuring high altitude N_2O , for example, from rocket-borne sampling devices, should be considered.

We thank Dr C. H. Jackman for his ion production rate calculation. The contributions of S.S.P. represent one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract NAS7-100, sponsored by NASA.

Received 3 December 1980; accepted 10 April 1981.

- Crutzen, P. J., Isaksen, I. S. A. & Reid, G. C. *Science* **189**, 457 (1975).
- Reid, G. C., Isaksen, I. S. A., Holzer, T. E. & Crutzen, P. J. *Nature* **259**, 177 (1976).
- Thorne, R. M. *Science* **195**, 287 (1977).
- Maloombe-Lewis, D. J. *Nature* **247**, 540 (1974).
- Zipf, E. C. *Nature* **287**, 523 (1980).
- Prasad, S. S. in *Solar-Terrestrial Influence on Weather and Climate* (eds McCormac, B. M. & Seliga, T. A.) 299-304 (Reidel, Boston, 1979).
- Zipf, E. C. & Prasad, S. S. *Nature* **287**, 525 (1980).
- Swider, W. *Geophys. Res. Lett.* **6**, 335 (1976).
- Cartwright, D. C., Chutjian, A. & Trajmar, S. *Phys. Rev.*, **A16**, 1041 (1977).
- Gilmore, F. *Pap. at Defense Nuclear Agency-Air Force Geophysical Laboratory Nuclear Weapons Effects Chemistry Conf.* Bedford (1980).
- Thorne, R. M. *Pure appl. Geophys.* **118**, 128 (1980).
- Imhof, W. L., Nightingale, R. W., Reagan, J. B. & Nakano, G. H. *J. atmos. terr. Phys.* **42**, 443 (1980).
- Turco, R. P. & Whitten, R. C. *NASA tech. Rep.* 1002 (NASA, Washington DC, 1977).
- Shimazaki, T., Ogawa, T. & Farrel, B. C. *NASA tech. Rep.* TND-8399 (NASA, Washington DC, 1977).
- Logan, J. A., Prather, M., Wofsy, S. & McElroy, M. B. *Phil. Trans. R. Soc. A* **290**, 187 (1978).
- Nicolet, M. & Cieslik, S. *Planet. Space Sci.* **28**, 105 (1980).
- Drummond, J. R. *et al. The Nimbus 7 User's Guide* (ed. Madrid, C. R.) (Goddard Space Flight Center, Greenbelt).

Field and structure behaviour of electrolytes

James W. Cobble & Richard C. Murray Jr

Department of Chemistry, San Diego State University, San Diego, California 92182, USA

Utpal Sen

Central Electrochemical Research Institute, Karaikudi 623006, Tamil Nadu, India

Measurements of the thermodynamic properties of electrolytes in hot water have a widespread practical and theoretical interest. The chemistry of many solutes changes rapidly as a solution is heated but the reasons for such changes have not been understood in terms of electrolyte theory. The temperature behaviour of the equilibrium thermodynamic properties is best seen by examining the ionic heat capacities of representative electrolytes, and we report here measurements just completed of $\bar{C}_{p2}^\circ$ of NaCl(aq) and BaCl₂(aq) up to 300 °C. The values are so negative that we now believe that the properties of electrolytes at higher temperatures are determined primarily by the electrostatic field of the ions and the bulk dielectric constant of water, and not by structural arrangements of ions and water molecules which characterize solutions at lower temperatures. These new phenomena are summarized in Fig. 1, which not only illustrates the large values at 300 °C for $\bar{C}_{p2}^\circ$ for NaCl(aq) and BaCl₂(aq) of -339 and -1,011 cal mol⁻¹ deg⁻¹, but also indicates a continued rapid change above 300 °C. The observations also suggest a need for reexamination of the theory of aqueous electrolytes at 25 °C, a temperature in the transition region between the two types of behaviour.

The standard state heat capacities for the electrolytes were obtained in a pressurized titanium-palladium calorimeter using the integral heat method¹ of dissolving salts to give very dilute solutions (down to 5×10^{-5} molal) and extrapolating the resulting heats of solution to infinite dilution. These heats at 10 temperatures between 50 and 300 °C were corrected to a pressure of 1 atm for calculation, using the molal volume data of Ellis and McFadden² up to 200 °C and extrapolated values up to 300 °C. The corrected data were fitted to a five-term power

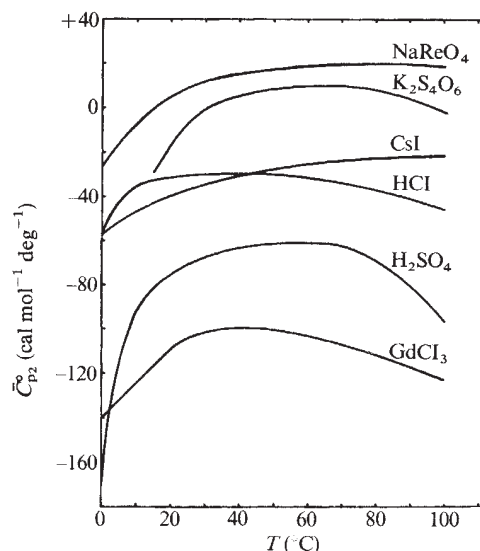


Fig. 1 Standard-state partial molal heat capacity of *a*, NaCl, and *b*, BaCl₂, 0-300 °C (solid heavy line) at P_{sat} . The Born electrostatic heat capacity function is indicated by the band.

Table 1 Comparison of 'field' and 'structure' electrolyte behaviour

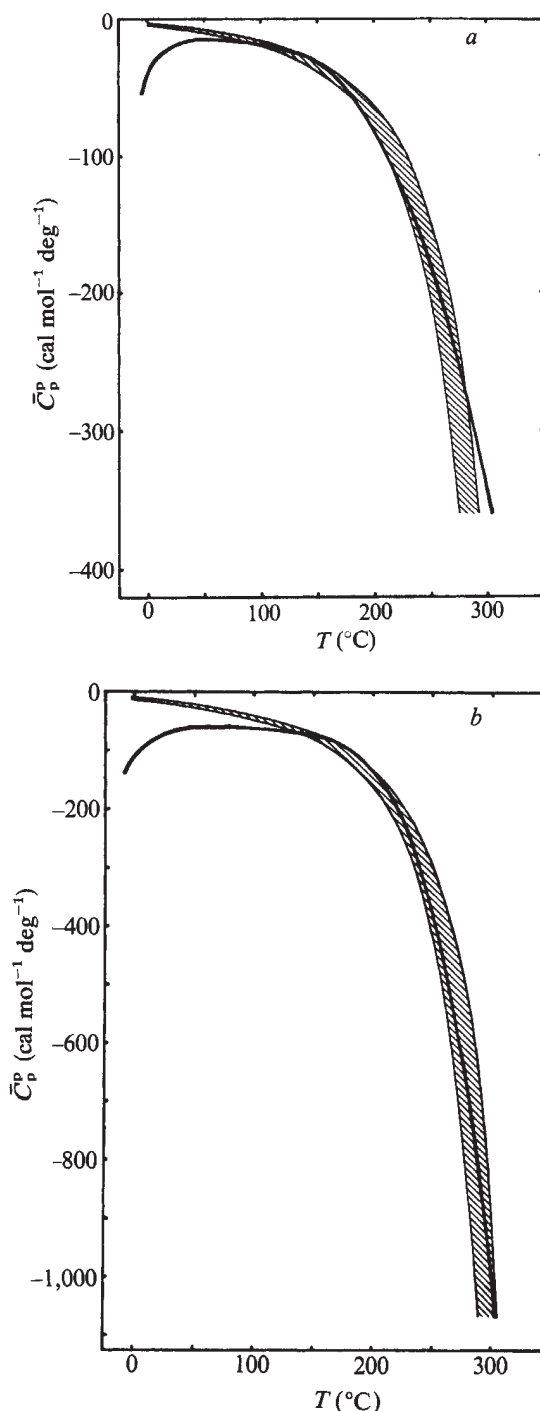
Structure	Field
1. $(d\bar{C}_{p2}^{\circ}/dT) > 0$	$(d\bar{C}_{p2}^{\circ}/dT) < 0$
2. Primary and secondary hydration of ions	Continuum of hydration around ions
3. Properties affected by structure of solvent	Properties dependent only on field of ion and D of solvent
4. Unequal hydration of cations and anions	
5. Structure-making and structure-breaking ions	

series and differentiated to obtain ΔC_p ; $\bar{C}_{p2}^{\circ}$ was then obtained by addition of the known heat capacities of the solids³ over the same temperature range. The 1-atm electrolyte heat capacities were then corrected back to the solvent pressure at each temperature (Fig. 1). The pressure corrections are negligible below 200 °C and were <4% even at the highest temperatures. The heat capacities below 100 °C for BaCl₂ and 200 °C for NaCl are from previously published data^{1,4}. The use of very dilute solutions is important as ionic strength effects are much larger at elevated temperatures. For example, the heat capacity of dilution of a 0.006 molal BaCl₂ solution to infinite dilution from the present data is roughly $-250 \text{ cal mol}^{-1} \text{ deg}^{-1}$ at 300 °C. The data obtained by flow calorimetry, which is limited to higher concentrations ($>0.1m$), also illustrates this point. Thus, Wood and Smith-Magowan⁵ indicate a heat capacity of dilution of NaCl from $m = 0.2433$ to $m = 0.0996$ of approximately $-100 \text{ cal mol}^{-1} \text{ deg}^{-1}$, at 175 atm. The heat capacity function, although difficult to obtain, is more temperature sensitive than other molal thermodynamic functions which are obtained by integration of $\bar{C}_{p2}^{\circ}$, a process which conceals many of the details of ion-solvent interactions.

Most investigators believe that the properties of water⁶⁻¹² and aqueous solutions^{6,12-16} near room temperature depend heavily on the hydrogen-bonded structure of the solvent. Competition between an ion and the normal solvent structure for solvent molecules led Frank and Wen¹³ to postulate three regions of solvation surrounding a central ion: the nearest-neighbour hydration sphere consisting of water molecules which are mostly, if not completely¹⁶, oriented towards the centre of the ionic charge; a second, transitional zone of non-structured water molecules; and finally, the outer zone comprised of the normal, structured 'bulk' water. However, on heating to higher temperatures, the hydrogen-bonded structure of bulk water begins to break down¹²⁻¹⁷ and the dielectric constant, D , falls from 87.85 at 0 °C to 20.05 at 300 °C, although the (DT) product is still large enough to sustain ionic behaviour⁸. The effects of the field of an ion, which, at 25 °C, are limited to small distances as a result of the high solvent dielectric constant and the structure built around the ion, should extend much further out from the ion and into the bulk solvent at 300 °C. The very negative values of $\bar{C}_{p2}^{\circ}$ obviously reflect the loss of degrees of freedom of increasing numbers of water molecules at higher temperatures as the following estimate illustrates. If $\bar{C}_p^{\circ}(\text{Na}^+) \approx \bar{C}_p^{\circ}(\text{Cl}^-) \approx -170$, then $\bar{C}_p^{\circ}(\text{Ba}^{2+}) \approx -670 \text{ cal mol}^{-1} \text{ deg}^{-1}$. Estimating that each water molecule involved in hydration loses, on the average, about one-quarter of its normal heat capacity (in tightly bound crystalline salts the loss is about one-half), we can easily calculate that Ba²⁺ influences $670/6.18$, or 109, water molecules at 300 °C; at 25 °C only 12 molecules are influenced. At 300 °C, the bulk density of water indicates an effective radius of 2.1 Å for a water molecule. Assuming a primary coordination number of six, the distance of influence extending out from the centre of Ba²⁺ corresponding to 109 water molecules is 13 Å (and similarly, 9 Å for Na⁺).

The low-temperature parts of the heat capacity curves in Fig. 1 indicate an opposite behaviour, where $\bar{C}_{p2}^{\circ}$ values become more negative with decreasing temperature, undoubtedly due to increasing solvent structural effects around the ion.

We propose the terms 'structure' (dominated) and 'field' (dominated) electrolytes to distinguish between these extremes in behaviour; in the former the structural effects of the solvent and its interaction with the ion are very important. In unshielded, field electrolytes, the solvent merely acts as a dielectric medium. In Fig. 1 the shaded zones indicate, within the uncertainty of the dielectric constant data,¹⁹ the Born^{20,21} electrostatic heat capacities for the two electrolytes, which have been obtained using the higher-temperature data to fix the radius terms; the agreement with simple electrostatic behaviour, above ~ 100 °C, is good. The temperature region where $(d\bar{C}_{p2}^{\circ}/dT) = 0$ is a good working definition of the transition region between the two types of behaviour, which is between 40 ° and 70 °C for many species (Fig. 2). Table 1 summarizes

**Fig. 2** Standard-state partial molal heat capacities for selected electrolytes^{29,30}, 0–100 °C.

some of the expected differences in behaviour between 'field' and 'structure' electrolytes.

There are other data which support these arguments. Twice differentiated free energies from precise equilibrium measurements in dilute solutions can give approximate values of $\bar{C}_{p2}^\circ$ at higher temperatures. The resulting large negative values, in $\text{cal mol}^{-1} \text{deg}^{-1}$ at 300°C , are -214 for $\text{H}^+ + \text{OH}^-$ from $K_w(\text{H}_2\text{O})$ (ref. 22), -210 for $\text{NH}_4^+ + \text{OH}^-$ from K_a (ref. 23) and the aqueous solubility²⁴ of ammonia, and -447 for $2\text{H}^+ + \text{CO}_3^{2-}$ from the aqueous dissociation^{25,26} and the solubility²⁷ of CO_2 . Furthermore, the large, negative partial molal volumes ($\sim -10^4 \text{ cm}^3$)²⁸ obtained for the ionic dissociation of $\text{NaCl}(\text{aq})$ and $\text{HCl}(\text{aq})$ above 360°C are believed to be consistent with our model. We are not aware of extensive heat capacity data in other solvent systems.

This paper is the result of work supported over previous years by the NSF and presently by the Electric Power Research Institute.

Received 26 November 1980; accepted 16 April 1981.

- Criss, C. M. & Cobble, J. W. *J. Am. chem. Soc.* **83**, 3223–3228 (1961).
- Ellis, A. J. & McFadden, I. M. *Geochim. cosmochim. Acta* **36**, 413–426 (1972).
- Kelley, K. K. *U.S. Bur. Mines Bull.* **584** (1960).
- Gardner, W. L., Mitchell, R. E. & Cobble, J. W. *J. phys. Chem.* **73**, 2025–2032 (1969).
- Wood, R. W. & Smith-Magowan, D. *Am. Chem. Soc. Ser. No. 133*, 569–580 (1980).
- Kavanau, J. Lee *Water and Solute–Water Interactions* (Holden-Day, San Francisco, 1964).
- Eisenberg, D. & Kauzmann, W. *The Structure and Properties of Water* (Oxford University Press, 1969).
- Narten, A. H. & Levy, H. A. *Science* **165**, 447–454 (1969).
- Frank, H. S. *Science* **169**, 635–641 (1970).
- Angell, C. A. *J. phys. Chem.* **75**, 3698–3705 (1971).
- Hagler, A. T., Scheraga, H. A. & Nemethy, G. *J. phys. Chem.* **76**, 3229–3243 (1972).
- Benson, S. W. *J. Am. chem. Soc.* **100**, 5640–5644 (1978).
- Frank, H. S. & Wen, W. Y. *Discuss. Faraday Soc.* **133**–140 (1957).
- Conway, B. E., Desnoyers, J. E. & Smith, A. C. *Phil. Trans. R. Soc. A* **256**, 389–437 (1964).
- Rossekinsky, D. R. *Chem. Rev.* **65**, 467–490 (1965).
- Bockris, J. O'M. & Saluja, P. P. S. *J. phys. Chem.* **76**, 2298–2310 (1972).
- Haggis, G. H., Hasted, J. B. & Buchanan, T. J. *J. chem. Phys.* **20**, 1452–1465 (1952).
- Cobble, J. W. *Science* **152**, 1479–1485 (1966).
- Bradley, D. H. & Pitzer, K. S. *J. phys. Chem.* **83**, 1599–1603 (1979).
- Born, M. *Z. Phys.* **1**, 45–48 (1920).
- Cobble, J. W. & Murray, R. C. *Jr Trans. Faraday Soc.* **64**, 144–149 (1978).
- Sweeton, F. H., Mesmer, R. E. & Baes, C. F. Jr *J. Solut. Chem.* **3**, 191–214 (1974).
- Hitch, B. F. & Mesmer, R. E. *J. Solut. Chem.* **5**, 667–679 (1976).
- Jones, M. E. *J. phys. Chem.* **67**, 1113–1115 (1963).
- Read, A. J. *J. Solut. Chem.* **4**, 53–70 (1975).
- Helgeson, H. A. *J. phys. Chem.* **71**, 3121–3136 (1967).
- Ellis, A. J. & Golding, R. M. *Am. J. Sci.* **261**, 47–60 (1963).
- Pearson, D., Copeland, C. S. & Benson, S. W. *J. Am. chem. Soc.* **85**, 1047–1049 (1963).
- Gardner, W. L., Jekel, E. C. & Cobble, J. W. *J. phys. Chem.* **73**, 2017–2020 (1969).
- Cobble, J. W., Stephens, H. P., McKinnon, I. R. & Westrum, E. F. Jr *Inorg. Chem.* **11**, 1669–1674 (1972).

Blast dynamics at Mount St Helens on 18 May 1980

Susan Werner Kieffer

US Geological Survey, Flagstaff, Arizona 86001, USA

At 8.32 a.m. on 18 May 1980, failure of the upper part of the north slope of Mount St Helens triggered a lateral eruption ('the blast') that devastated the conifer forests in a sector covering $\sim 500 \text{ km}^2$ north of the volcano. I present here a steady flow model for the blast dynamics and propose that through much of the devastated area the blast was a supersonic flow of a complex multiphase (solid, liquid, vapour) mixture. The shape of the blast zone; pressure, temperature, velocity (Mach number) and density distributions within the flow; positions of weak and strong internal shocks; and mass flux, energy flux, and total energy are calculated. The shape of blast zone was determined by the initial areal expansion from the reservoir, by internal expansion and compression waves (including shocks), and by the density of the expanding mixture. The pressure within the flow dropped rapidly away from the source of the blast until, at a distance of $\sim 11 \text{ km}$, the flow became underpressured relative to

the surrounding atmosphere. Weak shocks within the flow subparallel to the east and west margins coalesced at about this distance into a strong Mach disk shock, across which the flow velocities would have dropped from supersonic to subsonic as the pressure rose back towards ambient. The positions of the shocks may be reflected in differences in the patterns of felled trees. At the limits of the devastated area, the temperature had dropped only 20% from the reservoir temperature because the entrained solids thermally buffered the flow (the dynamic and thermodynamic effects of the admixture of the surrounding atmosphere and the uprooted forest and soils into the flow are not considered). The density of the flow decreased with distance until, at the limits of the blast zone, 20–25 km from the volcano, the density became comparable with that of the surrounding (dirty) atmosphere and the flow became buoyant and ramped up into the atmosphere. According to the model, the mass flux per unit area at the source was $0.6 \times 10^4 \text{ g s}^{-1} \text{ cm}^{-2}$ and the energy flux per unit area was 2.5 MW cm^{-2} . From the measured total ejected mass, $0.25 \times 10^{15} \text{ g}$, the total energy released during the eruption was 10^{24} erg or 24 megatons. The model, triggering of the eruption and the transition from unsteady to steady flow, and applications to eyewitness observations and atmospheric effects are discussed in ref. 1.

The eruption began through a vent centred between $\sim 1,900$ and $2,500 \text{ m}$ elevation $\sim 1 \text{ km}$ north of the old summit of the mountain. The 'devastated area' (the area in which trees have been destroyed) is a sector generally extending to the north of an east–west line through the vent formed by the landslide that triggered the event (Fig. 1). The devastated area is surrounded by a zone, the 'singled zone', in which trees are still standing, but on which remaining needles were singed.

Preliminary mapping of the direction of fall of the trees shows two major irregularly shaped zones: (1) an inner zone, the 'direct blast zone', in which the flow was approximately radial from the volcano and was relatively undeflected, in plan view, by even large topographic features; and (2) an outer zone, 'the channelized blast zone', in which the flow followed or was deflected, in plan view, by the local topography (Fig. 1). Streamlines of the flow were thus not simple radii from the source.

The reservoir from which the material erupted is assumed to have been a disk-shaped volume oriented vertically on the scarp formed on the north side of the mountain by the landslide. It is taken to be just large enough to contain the mass ejected by the blast, $0.25 \times 10^{15} \text{ g}$ (ref. 2). At an assumed pre-eruption bulk density of 2.0 g cm^{-3} , the reservoir would have had a volume of 0.12 km^3 . The frontal surface of this disk (the vent) is assumed to have been a rectangle 1 km in east–west dimension and 0.25 km in height; the reservoir would therefore have been 0.5 km deep. The acceleration of the blast material from rest to the vent velocity discussed below is assumed to have taken place within the reservoir¹.

The erupted material is characterized in this model by an initial pressure, P_0 , an initial temperature, T_0 , and an initial mass ratio, m , of solid plus liquid phases to the vapour phase. For an adiabatic, quasi-isentropic expansion, a mixture of vapour and fine-grained solid (or inert liquid) particles can be modelled as a thermally-equilibrated pseudo-gas and can be characterized by perfect gas equations of state³:

$$PV = R_b T \quad (1a)$$

$$PV^{\gamma_b} = \text{constant} \quad (1b)$$

where P is the pressure, V is the volume, γ_b is the isentropic exponent, and R_b is the gas constant of the blast mixture. Because of the fine-grained nature of the entrained solid particles, thermal transfer of heat to the gas phase from the solids is important; in such a case, the values of γ_b and R_b are

$$\gamma_b = \frac{C_{Pg} + mC_s}{C_{Vg} + mC_s} \quad (2a)$$

$$R_b = R_g / (1 + m) \quad (2b)$$

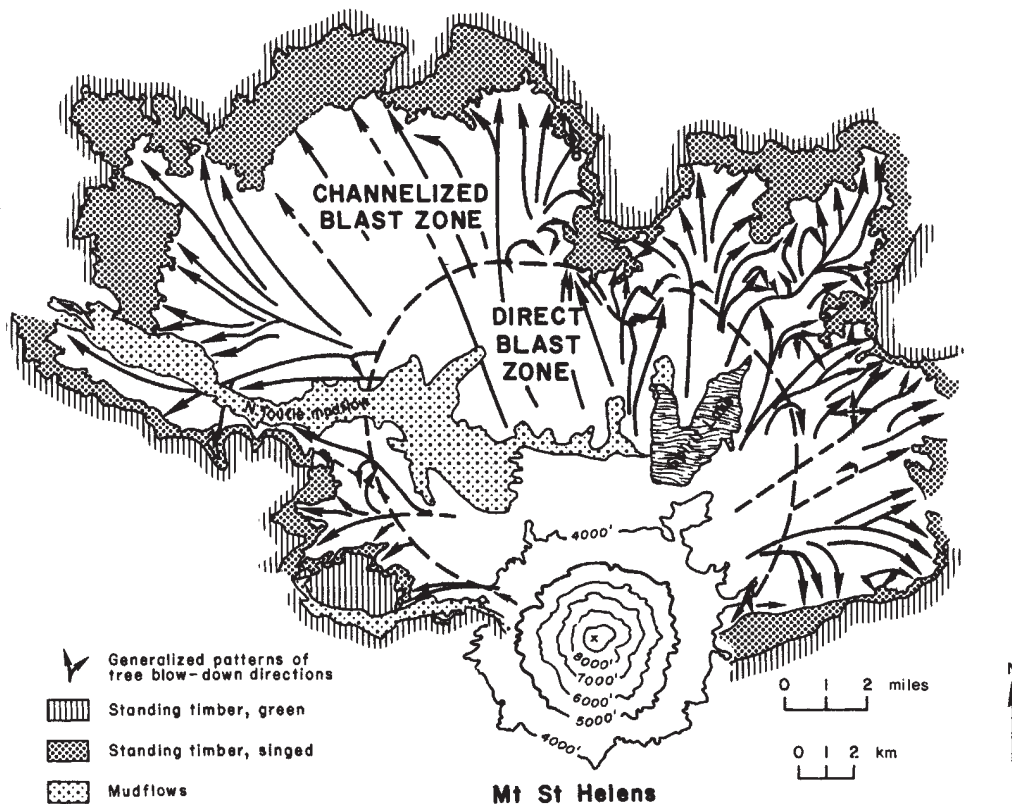
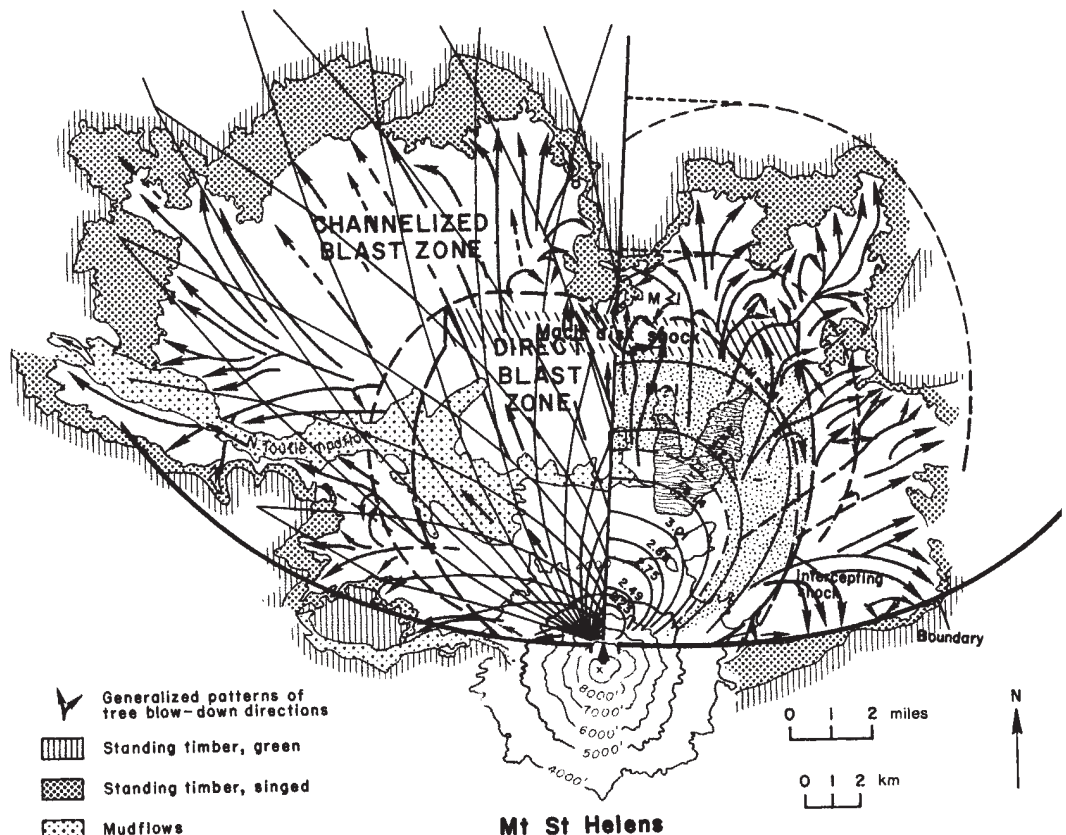


Fig. 1 The devastated area, and flow streamlines (arrowed vectors) as indicated by the alignment of fallen trees. The landslide at 8.32 a.m. initially broke ~1 km north of the old summit and formed the near-vertical vent from which the blast emanated. Relative to this vent position, the margin of the flow on the volcano bends south of due east and due west and then sweeps north-east and north-west as it diverges from the summit of the mountain. The singed zone is narrow compared with the size of the devastated area and streamlines of the flow are generally perpendicular to the margin. The boundary between the direct and channelized blast zones is shown by the short broken lines.

Fig. 2 The model flow field superimposed on the map of Fig. 1. The vent width has been taken as 1 km, has been placed at the 8,000-ft contour level north of the old summit of the mountain, and has been orientated 5° east of due north. As the model is symmetric about the axis of the vent it can be split into two halves. On the left, the mathematical characteristics of the solution are shown as thin lines radiating from the corner of the vent and from the mirror plane on the axis of symmetry. The arrows show the directions of the flow and would be the general direction of felled trees. The heavy line indicates the boundary of the flow, a 0.87-bar isobar in the model. The intercepting shock formed by reflection of the expansion waves from this boundary is shown as the long-dashed line connecting the Mach disk shock. Lines of constant Mach number and, therefore, constant pressure, temperature and density are shown. The approximate location of the Mach disk shock is shown by the sloping pattern 8 km across. Because estimates of flow density are needed for hazards zone prediction, the analytical solution is extrapolated into the subsonic region beyond the Mach disk shock (dashed contours), and is assumed to be about valid for density if the shocks are not strong. Each contour line is labelled by the value of the Mach number, M . The values of M , P/P_0 , T/T_0 and ρ/ρ_0 are, in the supersonic region, 2.23, 0.087, 0.91, 0.095; 2.49, 0.047, 0.89, 0.053; 2.75, 0.025, 0.87, 0.029; 2.88, 0.018, 0.86, 0.021; 3.01, 0.013, 0.85, 0.016; and 3.14, 0.009, 0.83, 0.011. Extrapolated into the subsonic region: 3.27, 0.006, 0.82, 0.007; 3.41, 0.004, 0.81, 0.006; 3.54, 0.003, 0.80, 0.004; 3.68, 0.002, 0.785, 0.002.



The approximate location of the Mach disk shock is shown by the sloping pattern 8 km across. Because estimates of flow density are needed for hazards zone prediction, the analytical solution is extrapolated into the subsonic region beyond the Mach disk shock (dashed contours), and is assumed to be about valid for density if the shocks are not strong. Each contour line is labelled by the value of the Mach number, M . The values of M , P/P_0 , T/T_0 and ρ/ρ_0 are, in the supersonic region, 2.23, 0.087, 0.91, 0.095; 2.49, 0.047, 0.89, 0.053; 2.75, 0.025, 0.87, 0.029; 2.88, 0.018, 0.86, 0.021; 3.01, 0.013, 0.85, 0.016; and 3.14, 0.009, 0.83, 0.011. Extrapolated into the subsonic region: 3.27, 0.006, 0.82, 0.007; 3.41, 0.004, 0.81, 0.006; 3.54, 0.003, 0.80, 0.004; 3.68, 0.002, 0.785, 0.002.

where C_{Pg} and C_{Vg} are, respectively, the specific heats at constant pressure and volume of the gas, and C_s is the specific heat of the solid phase (for which C_P and C_V are assumed to be approximately equal). Although the model can easily be scaled with dimensional, pressure and temperature ratios, a model case for a single set of plausible initial conditions at Mt St Helens is presented here. For this case, it was assumed that 10% by volume of the flank of the volcano that became the blast material was filled with liquid water that vaporized to give the gas during the eruption; for an average rock density of 2.5 g cm^{-3} and liquid water density of 1 g cm^{-3} , this assumption gives $m = 25$, $\gamma = 1.04$, and $R_b = 17.8 \times 10^4 \text{ erg C}^{-1} \text{ mol}^{-1}$. The initial source pressure is assumed to have been 12.5 MPa, appropriate to 650 m of overburden, and the initial temperature is assumed to have been 600 K, the saturation temperature of pure H_2O at 12.5 MPa. The initial density of a mixture of rock (assumed to have vesiculated or cracked to an average density of 2.0 g cm^{-3}) and steam (at a density of 0.074 g cm^{-3}) at 12.5 MPa pressure would have been 1.0 g cm^{-3} .

Because the reservoir was at such high pressure compared with the ambient atmospheric pressure, the flow would have choked at the vent and exited at the choked or sonic flow velocity. For steady flow, this velocity is

$$u_s^* = \left(\frac{2}{\gamma_b + 1} \right)^{1/2} c_{ob} \quad (3)$$

where c_{ob} is the speed of sound of the blast mixture at rest:

$$c_{ob} = (\gamma_b R_b T)^{1/2} \quad (4)$$

For the mixture postulated, c_{ob} is 105 m s^{-1} and the choked velocity is 104 m s^{-1} . The pressure at the vent would have decreased to $P^* = 0.60 P_0$ (7.5 MPa), the temperature to $T^* = 0.98 T_0$ (588 K), and the density to $\rho^* = 0.61 \rho_0$ (0.61 g cm^{-3}). The mass flux, controlled at the vent by choking, was

$$\dot{m}^* = \rho^* u^* A^* \quad (5)$$

where ρ^* , u^* and A^* are, respectively, the density, velocity and area at sonic conditions at the vent. With an average temperature 325 K above ambient, a specific heat of $0.8 \times 10^7 \text{ erg g}^{-1} \text{ K}^{-1}$ for the entrained solid or liquid phases, $4.2 \times 10^7 \text{ erg g}^{-1} \text{ K}^{-1}$ for the steam, and $2.3 \times 10^{10} \text{ erg g}^{-1}$ latent heat of vaporization of the steam, the thermal energy flux or power unit area was $2.5 \times 10^{13} \text{ erg s}^{-1} \text{ cm}^{-2}$, or 2.5 MW cm^{-2} . From the measured mass ejected and the assumed initial temperature of 600 K, the total thermal energy released on cooling to ambient temperature (273 K) was 10^{24} erg or 24 megatons, of which 7 megatons were released during the propagation of the blast through the devastated area, and 17 megatons during penecontemporaneous condensation of the steam and cooling of the condensed water and entrained solids.

After leaving the vent at sonic velocity, the material would have expanded supersonically in all directions into the lower-pressure atmosphere. In the absence of gravity, friction, phase changes, and addition or loss of mass from a flowing fluid, the steady-state equations of motion for an expanding fluid are:

$$\bar{\nabla} \cdot (\rho \bar{u}) = 0 \quad (\text{continuity}) \quad (6a)$$

$$\bar{\nabla} \times \bar{u} = 0 \quad (\text{irrotationality}) \quad (6b)$$

$$\bar{\nabla} \left(\frac{\bar{u}^2}{2} \right) + \frac{1}{\rho} \bar{\nabla} P = 0 \quad (\text{momentum}) \quad (6c)$$

Here ρ is the density, $\bar{u}$ is the vector velocity, P is the pressure, and $\bar{\nabla}$ is the spatial differential operator. These equations, supplemented by the equation of state of the fluid and the boundary condition that the edge of the flow is everywhere at atmospheric pressure, define the steady state condition of the fluid. They were solved by the method of characteristics within the region where shocks do not exist; various scaling relations

adopted from the aeronautics literature⁴ were used to estimate the position and size of shocks, and the shape of the flow boundary outside the regions of shocks.

The flow model for the boundary and reservoir conditions assumed above for Mt St Helens is shown in Fig. 2. The characteristics of the solution can be thought of as expansion waves across which the flow accelerated as the pressure, density and temperature dropped. For the model conditions, the flow would have expanded initially through an angle of 96° near the vent. At the boundary of the flow, the pressure was reduced to the local atmospheric pressure and the initial expansion waves were reflected back into the flow as compression waves that coalesced into the intercepting shocks shown near the margins of the flow. These reflections deflected the flow from its initial angle of 96° towards $\sim 60^\circ$ at a distance of 10 vent diameters out from the mountain.

As the flow diverged, the pressure dropped rapidly away from the vent. Isobars show that the pressure dropped below ambient outside a zone ~ 6 vent diameters from the vent (the negative excess pressure zone is shown as a stippled area in Fig. 2). This low-pressure core was stable within the flow because the compression waves could not be reflected back into the flow quickly enough to prevent overexpansion. The Mach disk shock is a consequence of the overexpansion and, across this shock, the pressure rose back towards ambient, and the flow velocities decreased from supersonic to subsonic. I proposed that the position of the Mach disk shock and the lateral intercepting shocks coincides with the boundary between the direct and channelized blast zones, and that the transition from supersonic to subsonic flow across this boundary accounts for differences in tree blowdown patterns.

As the pressure within the flow decreased, the density likewise decreased rapidly (although temperatures were remarkably uniform across the flow because of the buffering effect of the entrained solid and liquid phases which acted as a heat source for the expanding vapour). The density dropped from a value of 1.0 g cm^{-3} in the reservoir to a value very near that of the atmosphere (which is assumed to have had 2–3 times its normal density because of dust-ladenness during the unsteady development of the flow) at ~ 20 – 25 km distance from the vent. (It is assumed that the densities given by the solution interior to the shocks can be extrapolated across the shocks, a reasonable approximation if the shocks are not too strong.) Thus, as the density of the flow became equal to that of the (dirty) atmosphere, the flow became buoyant and simply lifted from the ground and ramped up into the atmosphere. The singed zone then is the area across which the flow was still close enough to the ground to cause thermal effects as it rose.

Numerical uncertainties for the calculated variables would be meaningless in view of the uncertainties in physical principles that we still face in interpreting the blast flow dynamics. The important reservoir parameters suggested by this model are the initial pressure (because it determines the amount of lateral divergence as the flow leaves the vent), the mass ratio, m , of solid-to-vapour components (because it determines the isentropic exponent, γ , and thus the thermodynamic history of the flow), and the vent diameter (because the dimensions of the devastated area scale with this dimension). From the model, however, it seems that, given reasonable *a priori* knowledge of the location of a potential vent and possible initial reservoir conditions, a hazards zone for blast-like events could be estimated by a relatively simple calculation. Models such as this and, as available, computer models that can calculate flow properties across internal shock waves, will be necessary for delineation of blast hazard zones because little evidence of the blast events is preserved in the geological record.

Received 9 February; accepted 25 March 1981.

1. Kieffer, S. W. *U.S. geol. Surv. Prof. Pap.* 1250 (in the press).
2. Moore, J. G. & Albee, W. C. *U.S. geol. Surv. Prof. Pap.* 1250 (in the press).
3. Wallis, G. B. *One-dimensional Two-phase Flow* (McGraw Hill, New York, 1972).
4. JANNAF (Joint Army, Navy, NASA, Air Force) *Handbook of Rocket Exhaust Plume Technology*, Ch. 2 (CPIA Publ. 263, 1975).

A model for the origin of the alkaline complexes of Egypt

Phil de Gruyter & Thomas A. Vogel

Department of Geology, Michigan State University, East Lansing, Michigan 48824, USA

The alkaline complexes in the Nubian Shield of Egypt range in age from 554 to 89 Myr BP (refs 1–3) and, where data are available, have similar and primitive initial strontium ratios of $^{87}\text{Sr}/^{86}\text{Sr}$ (ref. 1). These intrusive complexes are similar with respect to petrography, major element chemistry, rare earth element distribution and other trace element distribution^{1,4–7}; they are intimately associated with major lineaments in the Nubian Shield (Fig. 1)⁸. It is likely that all of them had a similar origin, not directly related to rifting and/or doming. We suggest here that the origin of these complexes is due to alkaline melts having been formed in the asthenosphere by shear heating, caused by changes in plate motion. These melts were emplaced along reactivated Pan-African fractures or pre-existing zones of weakness.

The major structural features of the Nubian Shield were formed during the Pan-African compression-accretion process (1,000–550 Myr BP)^{9–11}. This NE–SW oblique compression (see Fig. 4) produced both NE–SW and NW–SE zones of weakness^{9,12}. Most of the alkaline complexes of Egypt are associated with these lineaments.

The oldest known² alkaline intrusion in the Eastern Desert of Egypt is Wadi Dib, 554 Myr BP, emplaced during the last stages of the Pan-African event⁹. This complex is associated with the north-west trending Najd transcurrent fault system (Fig. 1) which has ~240 km of lateral displacement¹³ and developed at about the same time span (580–530 Myr BP)¹⁴ as the emplacement of the alkaline complex of Wadi Dib.

Zargat Naam (404 Myr BP)² may be associated with a north-west lineament and Tarbit North is located near the intersection of two lineaments (Fig. 1). El Gezira is dated² at 229 Myr BP and is associated with the same major north-east trending lineament as El Naga, El Mishbeh, Nigrub El Tahiani and Nigrub El Fogani. The ages² of these complexes are 145, 142, 140, 139, 132 Myr BP respectively. El Kahfa and Abu Khrug (89 and 91 Myr BP)¹ are associated with the same set of north-westerly trending lineaments (Fig. 1).

Although the alkaline complexes of the Egyptian Eastern Desert have similar overall chemical trends, there is a distinct change in the degree of alkalinity and silica saturation with age of the complexes⁴. The older complexes are generally silica oversaturated and relatively low in alkalis, whereas the younger complexes are dominated by silica-undersaturated rock types and are rich in alkalis (Fig. 2). This change in alkalinity may be a direct result of the cooling of the lithosphere after formation of the Pan-African thermal event. Higher temperatures after the Pan-African event would be conducive to larger degrees of melting at shallow levels, thus producing less alkaline magmas. As the lithosphere cooled, smaller degrees of melting occurred at greater depths, and more alkaline melts would be produced. Thus the chemical trends observed in the Egyptian alkaline complexes are consistent with decreasing thermal gradients in the Phanerozoic.

The six periods of alkaline magmatism in Egypt (554, 404, 351, 229, 145–132, 91–89 Myr BP) are all synchronous with known changes of plate motion. Changes in plate motion can be inferred from: (1) abrupt changes in global sea-level curves¹⁵; (2) the polar wandering curve for Africa since Pan-African time; and (3) the motion of adjacent plates during the past 180 Myr. Table 1 summarizes these data.

In comparison with tholeiitic magmatism, it is widely accepted that alkaline magmatism has a relatively deep-seated origin^{16–19}

and may be associated with relatively small degrees of partial melting of an enriched mantle source^{20–22}. These melts ascend preferentially along zones of weakness.

In light of recent thermo-mechanical modelling^{23–30}, a shear heating model for alkaline melts is reasonable for the production of alkaline magmas in areas of relatively low heat flow. Briefly, thermo-mechanical modelling indicates that an increase in plate velocity or a rapid change in direction of motion, causes an increase of the shear stress on the base of the lithosphere. This leads to temperature increases, viscosity decreases, and further plate stress and velocity changes. In these conditions the asthenosphere, the zone of decoupling between the rigid lithosphere and the viscous mesosphere, may be the locus for metasomatic fluids. In this region melts may be generated in sufficient quantities to become a significant source of alkaline magmatism^{31–34}. This melt production probably occurs at different levels³⁶ in a potentially unstable, thermally-active 'boundary zone' of the lower lithosphere–asthenosphere³⁵ (Fig. 3). It is suggested here that alkaline complexes of the Egyptian Eastern Desert area may be directly related to acceleration in plate motion. In this model, melting results from the internal dissipation of heat produced by this motion. Factors which are likely to have a bearing on the incidence of alkaline magmatism are: the amount of thermal disequilibrium produced by localization of shear heating; the ease of ascent of the melts (presence of fractures); and the possible involvement of non-alkaline melts generated at shallower depths. This approach to alkaline magmatism takes into account several features common to many alkaline complexes, including those in Egypt. They are: (1) Alkaline magmatism is unrestricted in its geographic and tectonic setting. It occurs in every tectonic environment in the oceans or continents^{37–52}, and is often independent of any rifting and/or doming events. (2) Alkaline complexes are emplaced in terrains that have relatively low thermal gradients^{38,53}. (3) They

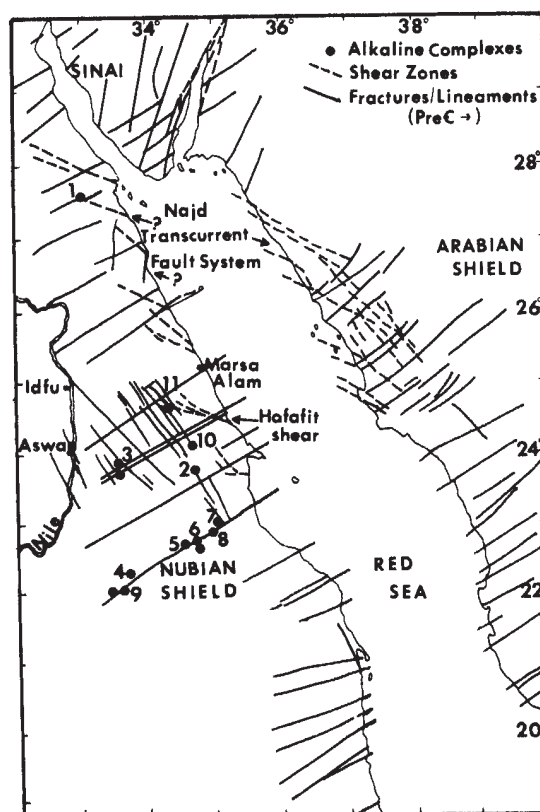


Fig. 1 1, Wadi Dib 554 Myr BP; 2, Zargat Naam 404 Myr BP; 3, Tarbit North 351 Myr BP; 4, El Gezira 229 Myr BP; 5, El Naga 145 Myr BP; 6, El Mishbeh 142 Myr BP; 7, Nigrub El Tahiani 140 Myr BP; 8, Nigrub El Fogani 139 Myr BP; 9, Mansouri 132 Myr BP; 10, El Kahfa 91 Myr BP; 11, Abu Khrug 89 Myr BP. Dates from refs 2, 3. Map and tectonic lineaments after ref. 8.

Table 1 Changes in plate motion

Periods of tectonic activity	Periods of rapid polar wandering (Africa) (Myr) ⁷⁰	Rapid changes in sea level ¹⁵ (Myr)	Periods of relative motion between Africa and Europe (Myr)	Periods of alkaline magmatism in Egypt ¹⁻³ (Myr)
Pan-African event of Eastern Desert, Egypt 1,000–550 Myr BP ⁹⁻¹¹ . Large-scale transcurrent faulting (Najd fault system, active 580–530 Myr BP towards the end of Pan African event or Hijaz Orogenic Cycle) ^{13,14,87}	680–660	480 410–390		554 404
Initial break-up of Pangea ⁷¹ leading to extension of North America–Africa/South America	400–220	350 330–320		351
Overall compression and subduction of Palaeo-Tethys north of Afro-Arabia; possible back-arc-type spreading of Cimmerian land mass from NE portion of Gondwana (including Afro-Arabia) and formation of Neo-Tethys ⁷²		280 220–185		229, 223, 221
Initial rifting of the South Atlantic: Africa–South America and Benue Trough extension		140	165–80	145, 142, 140
Corresponding compression in Afro-Arabia and Tethyan domain strike-slip faulting, closely followed by extension between Iran and Arabia ^{73,74} and the possible separation of the Anatolian plate from Afro-Arabia ⁷⁵			Mid- to late Cretaceous	139, 132
Major shift in spreading of the South Atlantic and associated second major episode of alkaline magmatism ^{56,57,60,61} ~80–90 Myr BP Large-scale strike-slip faulting in Afro-Arabia ^{5,76-81} intense compression in Iran and Turkey with associated ophiolites ^{73,74,82-84} emplacement of Oman ophiolites in NE Arabia ⁸⁵	110–40	95 60	80–40 Late Cretaceous Lower Eocene	91,89
Red Sea doming and extension	40–0 Lower Eocene to Recent	30		Recent alkali magmatism associated with Red Sea opening

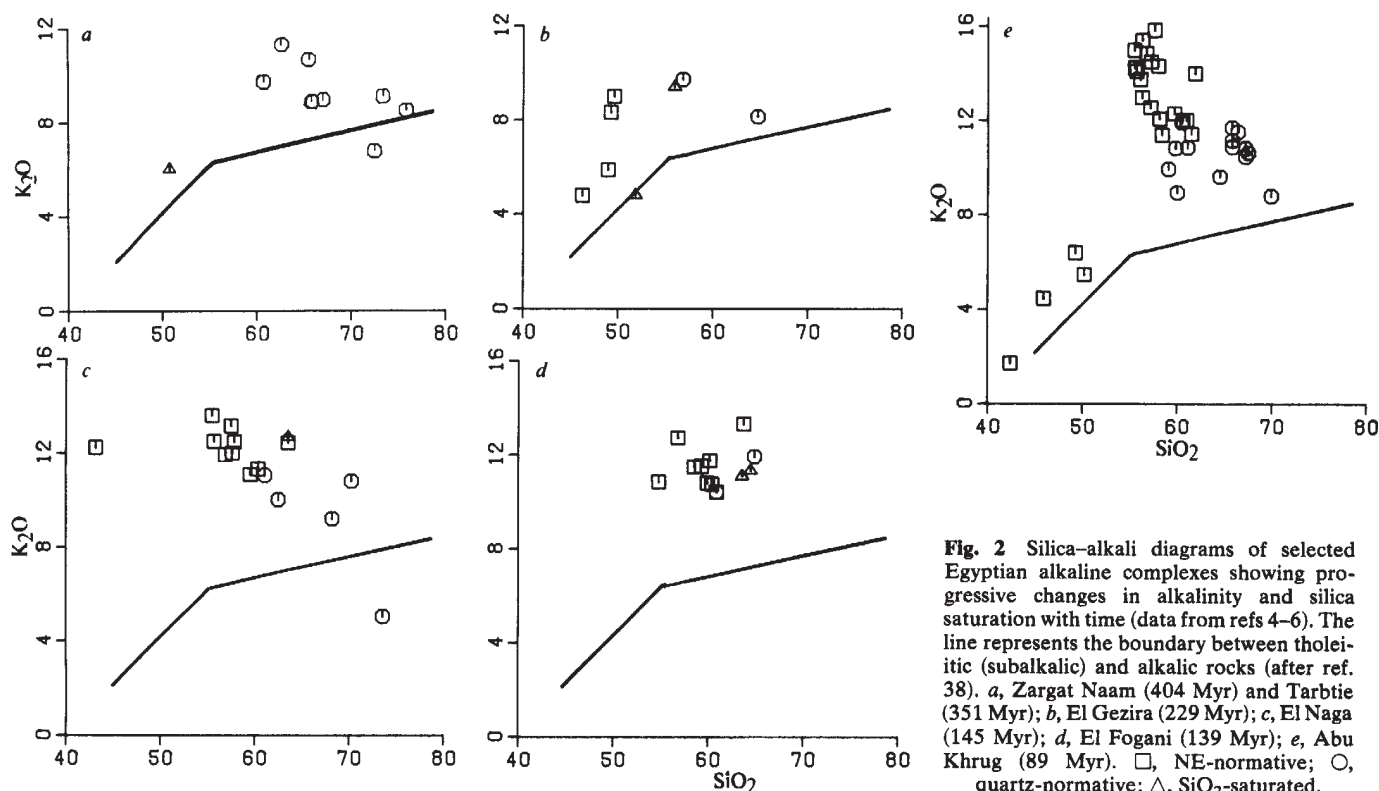
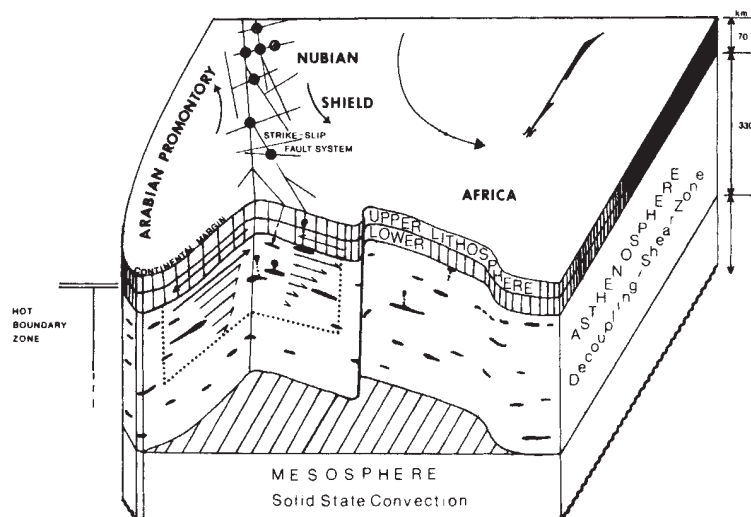


Fig. 2 Silica-alkali diagrams of selected Egyptian alkaline complexes showing progressive changes in alkalinity and silica saturation with time (data from refs 4–6). The line represents the boundary between tholeiitic (subalkalic) and alkalic rocks (after ref. 38). *a*, Zargat Naam (404 Myr) and Tarbtie (351 Myr); *b*, El Gezira (229 Myr); *c*, El Naga (145 Myr); *d*, El Fogani (139 Myr); *e*, Abu Khruh (89 Myr). □, NE-normative; ○, quartz-normative; △, SiO₂-saturated.

Fig. 3 Alkaline melt production through shear heating in the asthenosphere in response to major shifts in the motion of the African plate and their ascent along the reactivated strike-slip fault system (shear zones) between the Arabian Promontory and the Nubian shield. The generation of small melt fractions through internal heat dissipation is illustrated in the zone of decoupling between the lithosphere and mesosphere in response to major shifts in lithospheric motion. Similar small melt fractions may be produced in response to sustained opposing plate motion (for example, along a transform fault system). Dotted area outlines a general region of melt release and concentration of shear dissipation. No scale intended. ●, Alkaline magmatism.



have similar chemical trends and are enriched in some trace elements particularly the large ionic lithophile elements^{4,20,37,38,53,55}. (4) Egyptian alkaline complexes, along with other alkaline complexes, tend to occur during certain ages⁵⁶⁻⁵⁹, and these ages generally correspond to known accelerations in plate motion^{39,40,44,56,57,60-63}. (5) Many alkaline complexes are associated with long-lived zones of crustal weakness⁵⁷⁻⁶⁸.

The coincidence of alkaline magmatism in widely separated areas is well illustrated in the case of the Egyptian and Atlantic

alkaline provinces. The 224 Myr episode of alkaline magmatism in Egypt coincides with the first episode of alkaline magmatism in the White Mountain Magma Series⁶² along the crustal lineaments within the (still) adjoining North American continent. The 140 and 90 Myr episodes of alkaline magmatism in Egypt coincide with the two major episodes of alkaline magmatism occurring in the areas surrounding the South Atlantic. Alkaline magmatism around 140 Myr in Brazil, Uruguay, Bolivia, and western and southern Africa has been related to the initial rifting

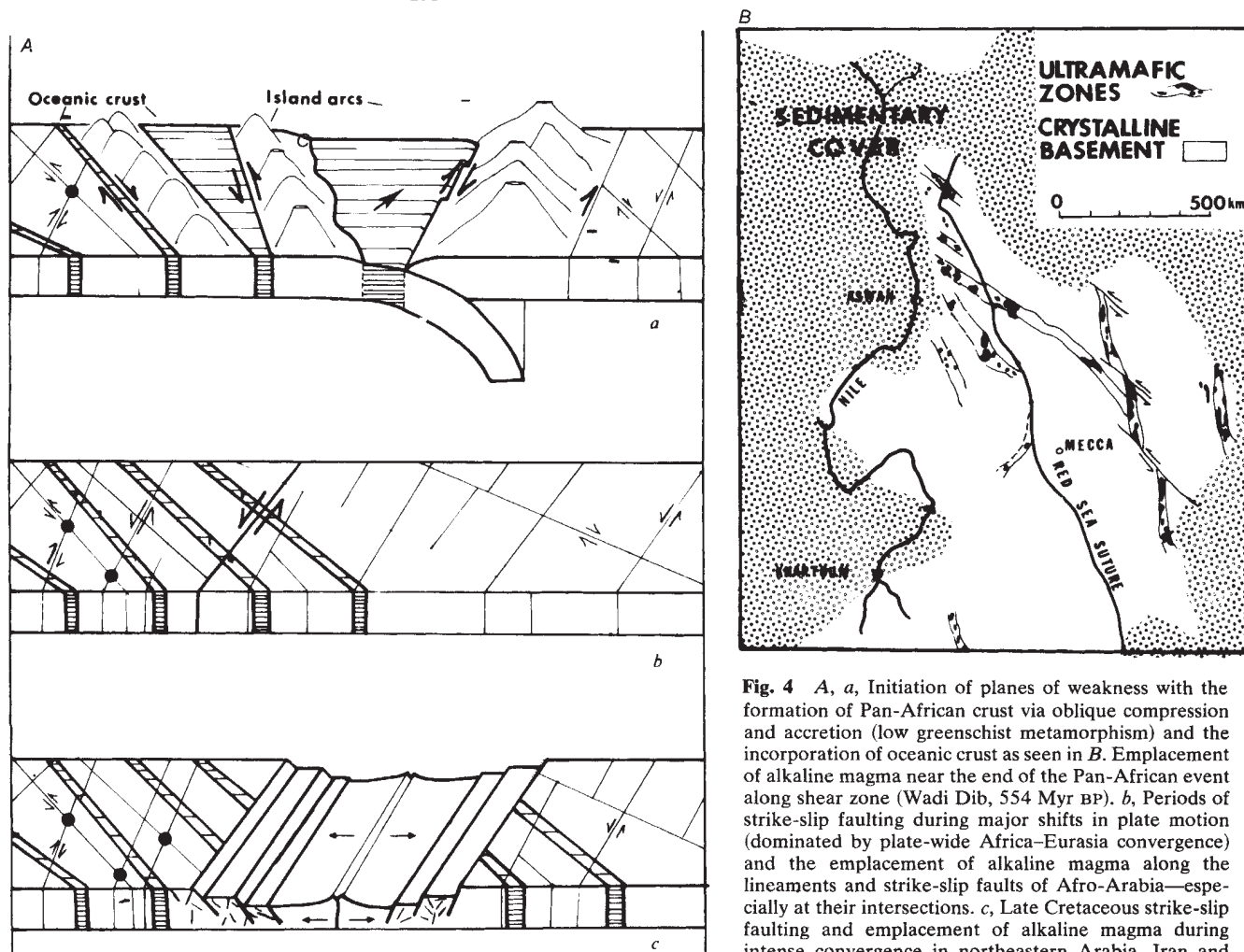


Fig. 4 A, a, Initiation of planes of weakness with the formation of Pan-African crust via oblique compression and accretion (low greenschist metamorphism) and the incorporation of oceanic crust as seen in B. Emplacement of alkaline magma near the end of the Pan-African event along shear zone (Wadi Dib, 554 Myr BP). b, Periods of strike-slip faulting during major shifts in plate motion (dominated by plate-wide Africa-Eurasia convergence) and the emplacement of alkaline magma along the lineaments and strike-slip faults of Afro-Arabia—especially at their intersections. c, Late Cretaceous strike-slip faulting and emplacement of alkaline magma during intense convergence in northeastern Arabia, Iran and Anatolia followed by mantle upwelling and rifting along

intensely weakened shear zones (Red Sea). B, Possible Pan-African ophiolites within Afro-Arabian shield with palinspastic reconstruction of Red Sea area, after ref. 12. ●, Alkaline magmatism.

of Africa from South America^{56-59,64}. The alkaline magmatism around 80–90 Myr has been correlated with a major shift in the spreading of the South Atlantic during the late Cretaceous^{56,57,60,63}. Considering the reciprocal plate motion involved in the opening of the South Atlantic and the closing of the Tethyan ocean basins⁶⁵⁻⁶⁷, the synchronous emplacement of these alkaline melts is not likely to be the result of unrelated mantle disturbances. We suggest that shear-heating in the asthenosphere due to these large changes in plate motion is a more likely explanation. Alternatively, these alkaline melts may be due to worldwide, synchronous, mantle disturbances⁶⁹.

The alkaline complexes of Egypt have been emplaced over a large time span (450 Myr) and seem to be associated with reactivated Precambrian lineaments. The ages of these alkaline complexes appear to be synchronous with changes in plate motion. Our suggestion for the origin of these alkaline complexes is that heat, necessary for melting, is produced through shear melting during acceleration of plate motion (Fig. 3) and that they are emplaced along pre-existing fractures or zones of weakness⁸⁶. This model may have general relevance to the origin of other alkaline magmas.

This research has been supported by NSF grant OIP 75-07943 administered by the University of South Carolina. We thank W. H. Kanes and E. M. El Shazly for continued interest and support, also Ahmed Abdallah for field assistance.

Received 20 January; accepted 2 April 1981.

- Lutz, T. M. thesis, Univ. Pennsylvania (1979).
- Serencsits, C. McC., El Ramly, M. F., Faul, H., Foland, K. A. & Hussein, A. A. *Geol. Surv. Egypt* (in the press).
- Hashad, A. H. & El Reedy, M. W. M. *5th Conf. on African Geology Abstr.*, Cairo, 44–45 (1979).
- deGruyter, P., & Vogel, T. A. (in preparation).
- El Ramly, M. F., Dereniuk, N. E., Budanov, V. I., Armonious, L. K. & Hayek, G. G. *Geol. Surv. Egypt Pap.* No. 51 (1969).
- El Ramly, M. F., Armanious, L. K., Budanov, V. I. & Deriuk, N. E. *Geol. Surv. Egypt Pap.* No. 52 (1969).
- El Ramly, M. F., Budanov, V. I. & Hussein, A. A. *Geol. Surv. Egypt Pap.* No. 53 (1971).
- Garson, M. S. & Krs, M. *Bull. geol. Soc. Am.* **87**, 169–181 (1976).
- Gass, I. G. *J. geol. Soc. Lond.* **134**, 129–138 (1977).
- Engel, A. E. J., Dixon, T. H. & Stern, R. *5th Conf. on African Geology Abstr.*, Cairo, 35–36 (1979).
- Engel, A. E. J., Dixon, T. H. & Stern, R. *J. Bull. geol. Soc. Am.* **91**, 699–706 (1980).
- Neary, C. R., Gass, I. G. & Cavanaugh, B. J. *Bull. geol. Soc. Am.* **87**, 1501–1512 (1976).
- Moore, J. Mc. *J. geol. Soc. Lond.* **136**, 441–454 (1979).
- Fleck, R. J. *et al. Bull. geol. Soc. Am.* **87**, 9–21 (1976).
- Vail, P. R., Mitchum, R. M. Jr & Thompson, S. III in *Seismic Stratigraphy—Applications to Hydrocarbon Exploration* Vol. 26, 83–98 (AAPG MEM, 1977).
- Green, D. H. & Ringwood, A. E. *Contr. Miner. Petrol.* **15**, 103–190 (1967).
- Bultitude, R. J. & Green, D. H. *J. Petrol.* **12**, 121–147 (1971).
- Kushiro, I., Syono, Y. & Akimoto, S. *J. geophys. Res.* **73**, 6023–6029 (1968).
- Green, D. H. *Phys. Earth planet. Inter.* **3**, 221–235 (1970).
- Frey, F. A., Green, D. H. & Roy, S. D. *J. Petrol.* **19**, 463–513 (1978).
- Gast, P. W. *Geochim. cosmochim. Acta* **32**, 1057–1086 (1968).
- Kay, R. W. & Gast, P. W. *J. Geol.* **81**, 653–682 (1973).
- Schubert, G., Froidevaux, C. & Yuen, D. A. *J. geophys. Res.* **81**, 3525–3540 (1976).
- Sammis, C. G., Smith, J. C., Schubert, G. & Yuen, D. A. *J. geophys. Res.* **82**, 3747–3761 (1977).
- Schubert, G. & Yuen, D. A. *Tectonophysics* **50**, 197–205 (1978).
- Yuen, D. A. & Schubert, G. *Geophys. Res. Lett.* **4**, 503–506 (1977).
- Yuen, D. A. & Schubert, G. *Geophys. J. R. astr. Soc.* **57**, 189–207 (1979).
- Froidevaux, C. & Souriau, M. *Phil. Trans. R. Soc. A288*, 387–392 (1978).
- Melosh, H. J. *Tectonophysics* **35**, 363–390 (1976).
- Melosh, H. J. & Ebel, J. *Geophys. J. R. astr. Soc.* **59**, 419–436 (1979).
- Shaw, H. R. *J. Petrol.* **10**, 510–535 (1969).
- Shaw, H. R. *Bull. geol. Soc. Am.* **84**, 1505–1526 (1973).
- Jackson, E. D., Shaw, H. R. & Bargar, K. E. *Earth planet. Sci. Lett.* **26**, 145–155 (1975).
- Jackson, E. D., Silver, E. A. & Dalrymple, G. B. *Bull. geol. Soc. Am.* **83**, 601–618 (1973).
- Shankland, T. J. *Oregon Dept Geol. Miner. Ind. Bull.* **96**, 215–218 (1977).
- Koide, H. & Bhattacharji *Oregon Dept Geol. Miner. Ind. Bull.* **96**, 263–280 (1977).
- Schwarzer, R. R. & Rogers, J. J. W. *Earth planet. Sci. Lett.* **23**, 286–296 (1974).
- Miyashiro, A. *Contr. Miner. Petrol.* **66**, 91–104 (1978).
- Hawkins, J. W. & Natland, J. H. *Earth planet. Sci. Lett.* **24**, 427–439 (1975).
- DeLong, S. E., Hodges, F. N. & Arculus, R. J. *J. Geol.* **83**, 721–736 (1975).
- Appleyard, E. C. *Lithos* **7**, 147–169 (1974).
- Sigurdsson, H., Tomblin, J. F., Brown, G. M., Holland, J. G. & Arculus, R. J. *Earth planet. Sci. Lett.* **18**, 285–295 (1973).
- Arculus, R. J. *Bull. geol. Soc. Am.* **87**, 612–624 (1976).
- Arculus, R. J., DeLong, S. E., Kay, R. W., Brooks, G. & Sun, S. S. *J. Geol.* **85**, 177–186 (1977).
- Price, R. C. & Taylor, S. R. *Contr. Miner. Petrol.* **40**, 195–205 (1973).
- Smith, I. E. M., Chappell, B. W., Ward, G. K. & Freeman, R. S. *Earth planet. Sci. Lett.* **37**, 230–236 (1977).
- Johnson, R. W., Wallace, D. A. & Ellis, D. J. *Volcanism in Australia* (Elsevier, Amsterdam, 1976).
- Johnson, R. W., MacKenzie, D. E. & Smith, I. E. M. *Tectonophysics* **46**, 197–216 (1978).
- Robin, C. & Tournon, J. *Can. J. Earth Sci.* **15**, 1633–1641 (1978).
- Bevier, M. L., Armstrong, R. L. & Souther, J. G. *Geology* **7**, 389–392 (1979).
- Heming, R. F. *Lithos* **12**, 173–186 (1979).
- van Breeman, O., Aftalion, M. & Johnson, M. R. W. *J. geol. Soc. Lond.* **136**, 489–495 (1979).
- Barth, T. F. W. *Theoretical Petrology* (Wiley, New York, 1962).
- McBirney, A. R., Aoki, K. *Geol. Soc. Am. Mem.* No. **116**, 523–556 (1968).
- Baker, B. H., Goles, G. G., Leeman, W. P. & Lindstrom, M. M. *Contrib. Miner. Petrol.* **64**, 303–332 (1977).
- Marsh, J. S. *Earth planet. Sci. Lett.* **18**, 317–323 (1973).
- Sykes, L. R. *Rev. Geophys. Space Phys.* **16**, 621–688 (1978).
- Neill, W. M. *Nature phys. Sci.* **245**, 104–107 (1973).
- Darbyshire, D. P. F. & Fletcher, C. J. N. *Geology* **7**, 545–548 (1979).
- Opdyke, N. & MacDonald, W. D. *Earth planet. Sci. Lett.* **18**, 37–44 (1973).
- van Breemen, O. & Bowden, P. *Nature phys. Sci.* **242**, 9–11 (1973).
- Foland, K. A. & Faul, H. *Am. J. Sci.* **77**, 888–904 (1977).
- Herz, N. *Bull. geol. Soc. Am.* **88**, 101–112 (1977).
- Amaral, G. *Geochim. cosmochim. Acta* **31**, 117–142 (1967).
- Dietz, R. S. & Holden, J. C. *J. geophys. Res.* **75**, 4939–4956 (1970).
- Smith, A. G. *Bull. geol. Soc. Am.* **82**, 2039–2070 (1971).
- Hsu, K. J. *Nature* **233**, 44–48 (1971).
- Sorensen, H. (ed.) *The Alkaline Rocks* (Wiley-Interscience, New York, 1974).
- Vogt, P. R. *J. Geophys. Res.* **86**, 950–960 (1981).
- Briden, J. C. & Gass, I. G. *Nature* **248**, 650–653 (1974).
- Gass, I. G., Chapman, D. S., Pollack, H. N. & Thorpe, R. S. *Phil. Trans. R. Soc. A288*, 581–597 (1978).
- Sengor, A. M. C. *Nature* **279**, 590–593 (1978).
- Dewey, J. F., Pitman, W. C. III, Ryan, W. B. F. & Bonnin, J. *Bull. Geol. Soc. Am.* **84**, 3137–3180 (1973).
- Alavi, M. *Geology* **8**, 144–149 (1980).
- Laubscher, H. & Bernoulli, D. in *The Ocean Basins and Margins* Vol. 4A (eds Nairn, A. E. M., Kanes, W. H. & Stehli, F. G.) 1–28 (Plenum, New York, 1977).
- Youssef, M. I. *Bull. Am. Ass. Petrol. Geol.* **52**, 601–614 (1968).
- Swartz, D. H. & Arden, D. D. *Bull. Am. Ass. Petrol. Geol.* **44**, 1621–1637 (1960).
- Burek, P. J. in *Hot Brines and Recent Heavy Metal Deposits in the Red Sea* (eds Degens, E. T. & Ross, D. A.) 59–70 (Springer, New York, 1969).
- Abdel-Gawad, M. in *Hot Brines and Recent Heavy Metal Deposits in the Red Sea* (eds Degens, E. T. & Ross, D. A.) 25–37 (Springer, New York, 1969).
- Hancock, P. L. & Kadhi, A. *J. geol. Soc. Lond.* **135**, 339–347 (1978).
- Brown, G. F. *Phil. Trans. R. Soc. A267*, 75–87 (1970).
- Hsu, K. J. in *The Ocean Basins and Margins* Vol. 4A (eds Nairn, A. E. M., Kanes, W. H. & Stehli, F. G.) 29–75 (Plenum, New York, 1977).
- Elasser, W. M. *J. geophys. Res.* **76**, 1101–1112 (1971).
- Pinar-Erdem, N. & Ilhan, E. in *The Ocean Basins and Margins* Vol. 4A (eds Nairn, A. E. M., Kanes, W. H. & Stehli, F. G.) 277–318 (Plenum, New York, 1977).
- Glennie, K. W. *et al. Bull. Am. Ass. Petrol. Geol.* **57**, 5–27 (1973).
- Turcotte, D. L. & Ahern, J. L. *J. geophys. Res.* **83**, 767–772 (1978).
- Brown, G. F. *Tectonic Map of the Arabian Peninsula*, Map. AP-2 (Directorate General of Mineral Resources, Jiddah, 1972).

Archaeological evidence for meat-eating by Plio-Pleistocene hominids from Koobi Fora and Olduvai Gorge

Henry T. Bunn

Department of Anthropology, University of California, Berkeley, California 94720, USA

The importance of meat-eating in human evolution has long been a controversial subject¹⁻⁴. The best available evidence of hominid activities between 2 and 1.5 Myr ago is the archaeological record from two East African localities, Olduvai Gorge⁵, Tanzania, and Koobi Fora⁶, Kenya, which consists of scattered stone artefacts and fragmentary animal bones. The question^{7,8} of functional association between juxtaposed artefacts and bones would be largely settled if hominid-induced modifications were present on some bones. Comparative analyses of archaeological bone assemblages from Olduvai Gorge and Koobi Fora and of various modern bone assemblages with known taphonomic histories reveal direct evidence of early hominid butchering and marrow-processing activities.

The surfaces and broken edges of bones from Olduvai and Koobi Fora were examined macroscopically in a strong light. Three general categories of modification were found: (1) hominid-induced damage features, including butchery (cut) marks and hammer-related fracture patterns; (2) carnivore- and rodent-induced damage features, including gnaw marks and tooth-related fracture patterns; (3) weathering and post-depositional alterations, including pitting, grooving and cracking.

A series of very fine linear grooves on bone surfaces constitutes the most unequivocal evidence of hominid involvement with the bones. These grooves are cut marks made by hominids using knife-like stone flakes to remove skin from carcasses, separate articulated bones and detach meat adhering to bones. The grooves occur singly or in multiple sets and vary in length from several millimetres to several centimetres. Typically they

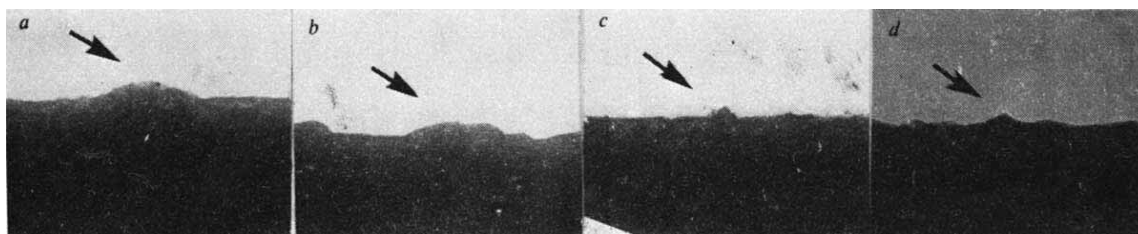


Fig. 1 High-resolution silicone impressions of grooved recent and fossil bone surfaces. Impressions were sectioned transverse to grooves and photographed; $\times 10.5$. Grooves on bone thus appear as positive bumps (arrows) on sectioned impressions of relatively flat bone surfaces. *a*, Broad, rounded cross-sectional shape of recent bone gnawed by carnivore. Width of groove ~ 1.42 mm. *b*, A series of overlapping small rodent gnaw marks on recent bone, showing flat to slightly rounded cross-sectional shape. Width of individual tooth mark ~ 0.71 mm. *c*, Much narrower V-shaped cross-section of cut mark made on recent bone using a flint flake. Width of cut mark ~ 0.36 mm. At FLK Zinj, mean width of 382 cut marks is 0.23 mm, and mean width of 146 carnivore gnaw marks is 0.70 mm. Typical carnivore gnaw mark is thus 3–4 times wider than a cut mark. *d*, Very narrow V-shaped cross-section of cut mark on 1.5-Myr-old fossil hippopotamus femur from site GaJi 5 at Koobi Fora.

are straight-sided and V-shaped in cross-section⁹, although some have a slightly flattened bottom, perhaps because of the dullness of the tool being used. These fine cut marks are readily distinguishable macroscopically from carnivore tooth marks, which are typically broader and rounded or U-shaped in section; from rodent gnaw marks, which are typically short, flat or round-bottomed, paired grooves; from irregular, dendritic pitting and grooving that results from the growth of plant roots against bone surfaces; and from faint, shallow linear grooves caused by contact with abrasive sediments (Fig. 1). Moreover, the cut marks have proved indistinguishable macroscopically from undoubted butchery marks, which are, of course, a common feature on archaeological bone debris from more recent periods. On replicas prepared for sectioning and microscopic scrutiny, and for scanning electron microscopy, the marks have also proved indistinguishable from experimentally induced cut marks. Methods for microscopically distinguishing cut marks from other kinds of damage have been developed by Potts and Shipman¹⁰.

Figure 2 shows a particularly well preserved example of this feature from site FxJj50 (ref. 11). The distal end of a very large, extinct bovid humerus has five separate, parallel grooves on its medial surface. Butchering experiments with fresh animal carcasses show that this is a prime location to cut when detaching the humerus from the radio-ulna. The remaining dozen or so examples of butchery marks from FxJj50 occur on limb and rib shafts where the likely aim was to detach meat. At FLK Zinj^{5,12}, with more complete skeletal representation and better bone

preservation, ~ 300 bones from equid, suid and all bovid genera present, retain cut marks.

Similar cut marks have been recognized on fossil bones which did not come from artefact-bearing archaeological sites. These were first recognized on specimens stored in palaeontological collections from Koobi Fora. Because of the high anthropological interest in such evidence, specific searches were made at Koobi Fora, and, in fact, bones with cut marks were recovered during the 1979 season from beds where hominid fossils occur but where artefacts and archaeological sites are sparsely

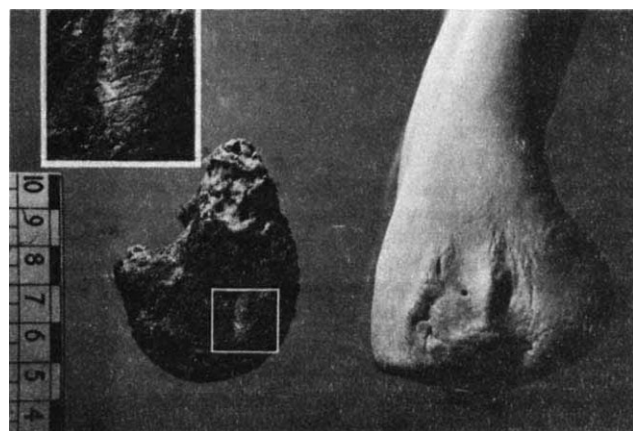


Fig. 2 Medial view of humerus of recent *Connocchaetes taurinus* (right) and distal end of humerus of *Megalotragus kattwinkeli* (left) from 1.5-Myr-old site FxJj 50 at Koobi Fora (scale in cm). Inset shows cut marks made by a hominid using a sharp-edged stone artefact; $\times 1.5$. Cut marks are present on 1.6% of FxJj 50 bone fragments and on 8.6% of the FLK Zinj bone fragments that are identifiable to skeletal part.

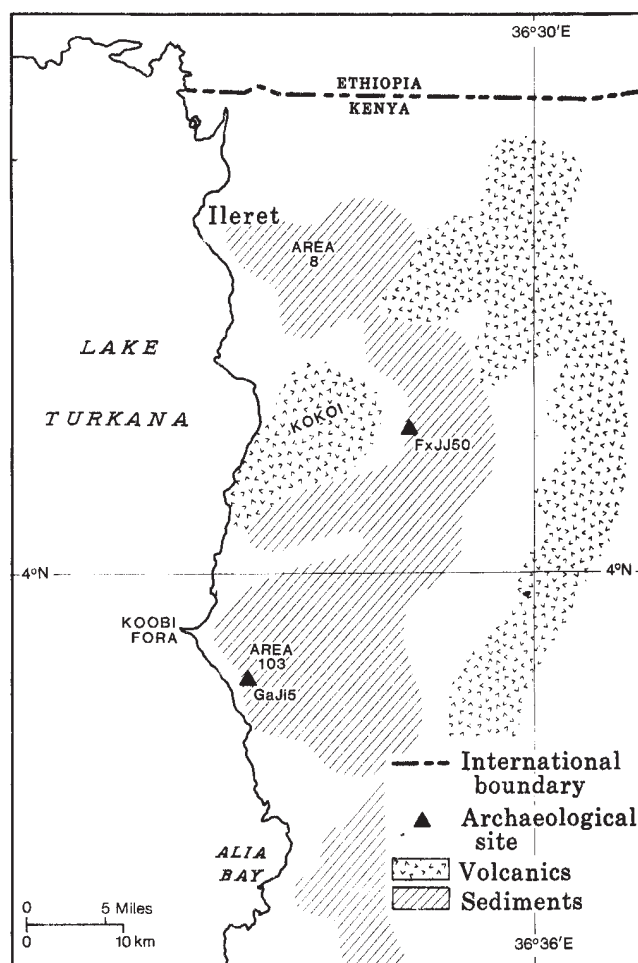


Fig. 3 Map of Koobi Fora study area east of Lake Turkana in northern Kenya, showing selected archaeological sites in relation to lake shore and to outcrops of volcanic and Plio-Pleistocene sedimentary rocks.

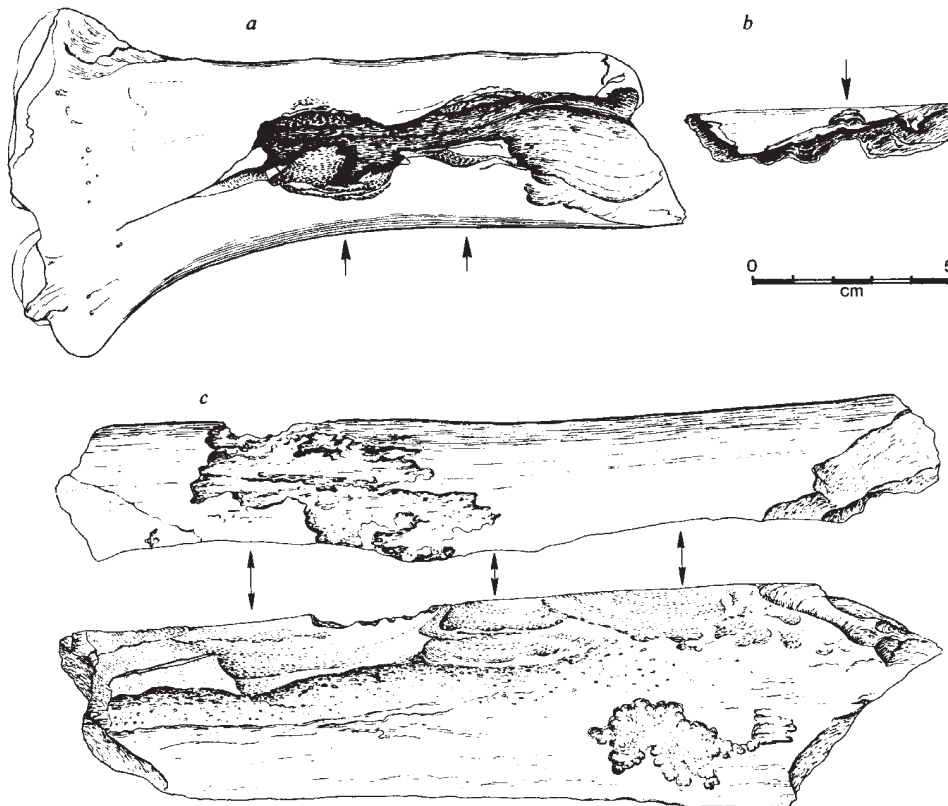


Fig. 4 Recent and fossil bones showing distinctive fracture patterns produced by hammer-stone percussion and carnivore chewing. *a*, Anterior view of proximal end of cow radius broken with hammer-stone. Repeated blows at two adjacent points (arrows) partially split the shaft and produced broad, arcuate indentations and broad bone flakes, several of which adhere to the bone shaft at points of impact. *b*, Recent limb shaft split by hyena chewing. Arrow indicates the point where the tooth contacted and split the shaft, producing relatively small, tooth-sized indentations. *c*, Two views of a very large fossil mammal radius shaft fragment from site FxJj 50: top view shows external surface and broken edge of bone; bottom view shows internal surface of bone with broad flake scars. Arrows mark points of percussion where broad, arcuate indentations indicate probable hammer-stone fracture by hominids (scale in cm). Drawn by B. Isaac.

represented, notably in exposures south of the Koobi Fora base camp (Area 103) and at Ileret (Area 8) (Fig. 3). These beds were deposited near the shoreline of an ancient lake. Source areas of stone raw material are located at least 10–15 km inland, as are almost all previously reported archaeological sites from the Koobi Fora Formation⁶. But at site GxJj5, in Area 103, a surface concentration of ~200 bone fragments but no stone artefacts includes 10 bones, with cut marks, from at least five different animals, including bovid, suid, giraffe and hippopotamus.

Distinguishing between bone assemblages in which a significant amount of fracture has been induced by hominids using hammers and assemblages in which the bones have been broken by chewing, trampling and other processes has proved complex^{11,13–19}. Hammer-stone blows produce broad internal flake scars on limb shaft pieces and bone flakes. Damage resulting from breakage by carnivores is similar, but the scars and flakes are much smaller. With carnivore chewing, the edges of cracks along which a bone splits open may again be indented and internally flaked at points where force has been applied, but instead of the broad, arcuate indentation produced by hammer-stone fracture, the indentation approximates the relatively small diameter of the carnivore tooth and thus typically shows a scalloped or denticulate edge with series of tooth-sized indentations along fractured edges (Fig. 4).

Figure 4 shows one example from the excavated assemblage at FxJj50 of a very large mammal radius shaft that was almost certainly broken by hominids using a hammer-stone. At FLK Zinj, the picture is more complex in that over 400 bones show damage attributable to chewing by medium-sized carnivores, and this damage potentially obscures signs of previous hammer-stone breakage by hominids. Although hominids may have broken some of the bones, none of over 50 bones flakes in the assemblage is large enough to preclude derivation from breakage by carnivores. The observations on bone fracture patterns, although less definite, do seem to be consistent with the cut mark evidence in implying significant hominid feeding on animal tissues.

The formation of these early archaeological sites involved a complex interplay of several key factors, including hominid butchering and marrow-processing activities, carnivore

scavenging and other animal activities, and various depositional and post-depositional processes. The excavated materials cannot therefore be viewed as a simple, complete record of hominid activities at the sites; rather, the bones and artefacts are a surviving sample of a dynamic system involving processes that added, modified and subtracted items of evidence¹¹.

The presence of cut marks and percussion marks on significant numbers of bones documents the involvement of early hominids nearly 2 Myr ago in cutting up animal carcasses and breaking open bones, presumably to obtain meat and marrow. In addition to indicating one use of stone artefacts, this direct evidence of early hominid diet allows us to dismiss models of human evolution which do not incorporate meat-eating as a significant component of early hominid behaviour. The documentation of meat-eating and the concentration of bones at particular places by early hominids lends strong support to the food-sharing model proposed by Isaac²⁰. The mode of acquisition of this meat, whether hunted, scavenged or both, remains uncertain and requires separate investigation.

These findings also provide a novel technique of mapping the distribution of hominid tool-using and feeding activities across ancient landscapes. In some areas where stone artefacts are too sparse to be detected, fossil bones with cut marks provide evidence of their former presence and usage. Searches for cut marks need to be made in areas that have yielded hominid and other fossil remains from sediments older than 2 Myr, in a time range that may pre-date the habitual discard of quantities of stone artefacts recognizable to archaeologists.

I thank J. D. Clark, G. Isaac, E. Kroll, R. Potts and P. Shipman for comments on an earlier draft of this paper. This research is part of the Koobi Fora Research Project, co-directed by G. Isaac and R. Leakey, of the National Museums of Kenya, and is supported by an NSF grant to G. Isaac.

Received 20 November 1980; accepted 18 March 1981.

1. Dart, R. A. *Transv. Mus. Mem.* **10**, 1–105 (1957).
2. Ardrey, R. *African Genesis* (Collins, London, 1961).
3. Ardrey, R. *The Hunting Hypothesis* (Atheneum, New York, 1976).
4. Zihlman, A. & Tanner, N. in *Female Hierarchies* (eds Tiger, L. & Fowler, H.) (Aldine, Chicago, 1978).

5. Leakey, M. D. *Olduvai Gorge Volume III: Excavations in Beds I and II, 1960–1963* (Cambridge University Press, 1971).
6. Isaac, G. L. & Harris, J. W. K. in *Koobi Fora Research Project Vol. 1* (eds Leakey, M. G. & Leakey, R. E.) (Clarendon, Oxford, 1978).
7. Binford, L. R. *J. anthropol. Res.* **33**, 493–502 (1977).
8. Isaac, G. L. *Phil. Trans. R. Soc. B* **292**, 177–188 (1981).
9. Walker, P. L. & Long, J. C. *Am. Antiq.* **42**, 605–616 (1977).
10. Potts, R. & Shipman, P. *Nature* **291**, 577–580 (1981).
11. Bunn, H. *et al. World Archaeology* **12**, 109–136 (1980).
12. Hay, R. L. *Geology of the Olduvai Gorge* (University of California Press, Berkeley, 1976).
13. Sutcliffe, A. J. *Nature* **227**, 1110–1113 (1970).
14. Brain, C. K. in *Human Origins: Louis Leakey and the East African Evidence* (eds Isaac, G. L. & McCown, E. R.) (Benjamin, Menlo Park, 1976).
15. Hill, A. P. in *Fossils in the Making: Vertebrate Taphonomy and Paleocology* (eds Behrens-meyer, A. K. & Hill, A. P.) (University of Chicago Press, 1980).
16. Sadek-Kooros, H. *Am. Antiq.* **37**, 369–382 (1972).
17. Myers, T. P. *et al. Am. Antiq.* **45**, 483–490 (1980).
18. Bonnichsen, R. in *Mammalian Osteo-Archaeology: North America* (by Gilbert, B. M.) (Missouri Archaeological Society, Columbia, 1973).
19. Noe-Nygaard, N. *Paleobiology* **3**, 218–237 (1977).
20. Isaac, G. L. *Scient. Am.* **238** (4), 90–108 (1978).

Cutmarks made by stone tools on bones from Olduvai Gorge, Tanzania

Richard Potts* & Pat Shipman†

*Department of Anthropology, Harvard University, Cambridge, Massachusetts 02138, USA

†Department of Earth and Planetary Sciences, The Johns Hopkins University, Baltimore, Maryland 21218, USA

Fossils from Olduvai Gorge, Tanzania, show cutmarks which establish that hominids were using stone tools on animal tissues during the Lower Pleistocene in Africa. We identified cutmarks by elimination of other likely causes of the marks on the bone surfaces, for example, gnawing or chewing by carnivores or rodents, and damage made by tools of excavators or preparators. This was achieved by comparing the marks on the fossils with those produced by known causes on modern bones, using scanning electron microscopy (SEM). Because the fossils occur as part of accumulations of animal remains in relatively undisturbed geological contexts, we conclude that there is a functional association between the stone artefacts and bones at these sites, rather than an accidental, postmortem association^{1,2}.

The problem in identifying cutmarks produced by stone edges is that similar grooves and marks can be made on bone surfaces by a variety of processes. As part of a study³ of the effects of various taphonomic processes on bones, the following types of mark were investigated: carnivore tooth scratches, incisal gnawing by rodents or carnivores, root etching, abrasion by windborne and waterborne sedimentary particles, and excavators' and preparators' marks. Each of these alterations to bone surfaces are macroscopically similar to cutmarks. Because the microscopic structure of a bone is a major determinant of its response to potentially destructive forces^{4–9}, several hundred such marks, produced on modern bones by known events, were inspected by SEM to clarify the interaction between these forces and the microscopic structure of the bone. This study yielded morphological criteria by which cutmarks of various types could be distinguished from the similar marks produced by other events or by normal anatomical features (Fig. 1). Samples are too small at present to develop quantitative criteria.

The controls and a sample of similar marks on fossil bones were inspected on epoxy replicas of bone surfaces; these replicas (which are expected to have a resolution of 0.25 µm) were coated with 200 Å of gold palladium to enhance SEM image quality¹⁰. Despite the low magnifications used in some cases, the depth of field and image quality of the SEM offered a substantial advantage over a light microscope.

Several types of cutmark were defined for the control sample, on the basis of their mode of production and microscopic characteristics. The term 'cutmarks' is used here as a general term for marks made by stone tools; microscopically, these are

elongated grooves with V-shaped cross-sections. Slicing marks are produced by drawing the edge of an artefact across a bone surface in a direction continuous with the long axis of the edge. This creates many fine, parallel striations, visible at a magnification of $\times 30$ – $\times 35$ or less, within each main groove because the edges of artefacts are not perfectly straight; each small deviation of the edge to one side or the other leaves its own microscopic track. Chopping marks are produced by striking the bone surface with an artefact at a roughly perpendicular angle—these have V-shaped cross-sections and small fragments of bone crushed inwards at the bottom of the main groove. Chopping marks are often broader at the top than slicing marks and do not show fine parallel striations. Scraping marks are formed by drawing an edge across a bone surface in a direction roughly perpendicular to the long axis of the edge. This action results in multiple, fine, parallel striations across a broad area of bone rather than confined to a single, elongated main groove (Fig. 1a–c).

No process has yet been discovered which produces marks that mimic slicing, chopping or scraping marks on a microscopic level. Tooth scratches and gnawing of carnivores produce grooves with rounded or flat bases respectively; both lack the fine parallel striations of slicing or scraping marks. Very fine tooth scratches may be as narrow as cutmarks and usually require magnification of $\times 20$ – $\times 50$ before they can be distinguished from cutmarks. Commonly, rodent gnawing marks are a series of short, very broad, parallel grooves which can be identified with the naked eye. Abrasion by windborne or waterborne sedimentary particles, or by collision of bones with rocks during hydraulic transport, removes external bone surfaces and opens up vascular canals originally lying below the surface, to give a grossly striated appearance. Microscopically, these canals differ from tooth scratches, in showing openings for even smaller vascular canals, and from slicing marks in lacking striations. Scratches produced on fossils with metal tools used for excavation or preparation are not always V-shaped in cross-section and do not show fine, parallel striations characteristic of cutmarks on fresh bones. The edges of such scratches or grooves are often much more irregular than tooth scratches or gnawing marks (Fig. 1d–g).

Numerous fossil bone specimens showing surface marks were found during a study of 12 excavated levels in Beds I and II, Olduvai Gorge¹¹. A sample of 75 bone surfaces, some with several marks, was chosen and replicated from fossils at each level. A wide range of skeletal parts, taxa and types of surface damage were included to make the sample as representative as possible. The surface of the fossils was brushed for several minutes with solvents before replication—this was necessary because preservatives completely obscure the microscopic characteristics. The replicas were inspected and photographed at magnifications comparable with those used for the control sample of modern bones.

In the replicated Olduvai sample ($N = 85$ marks on 75 bones), 24% of the specimens possessed marks which were considered to be cutmarks as they more closely resembled those produced by slicing, chopping or scraping than those produced by any other cause so far investigated. Isolated or multiple fine grooves, which could have been confused with cutmarks from a macroscopic inspection, often resembled tooth scratches or narrow gnawing marks much more closely in terms of microscopic detail. Table 1 gives the numbers of fossils with each type of mark; Fig. 2a shows a series of slicing and scraping marks on a fossil.

Significantly, several bones had both slicing marks and tooth scratches, including three specimens which showed intersections between these types of mark (Fig. 2b, c). Thus it is clear that, at least in some cases, hominids and carnivores were able to use the same parts of carcasses in turn. Furthermore, slicing marks are present on at least one specimen from FLK NN, Level 2, although no stone tools are known from that site. One possible interpretation of this evidence is that carnivores scavenged this bone from a hominid campsite and removed it to another area.

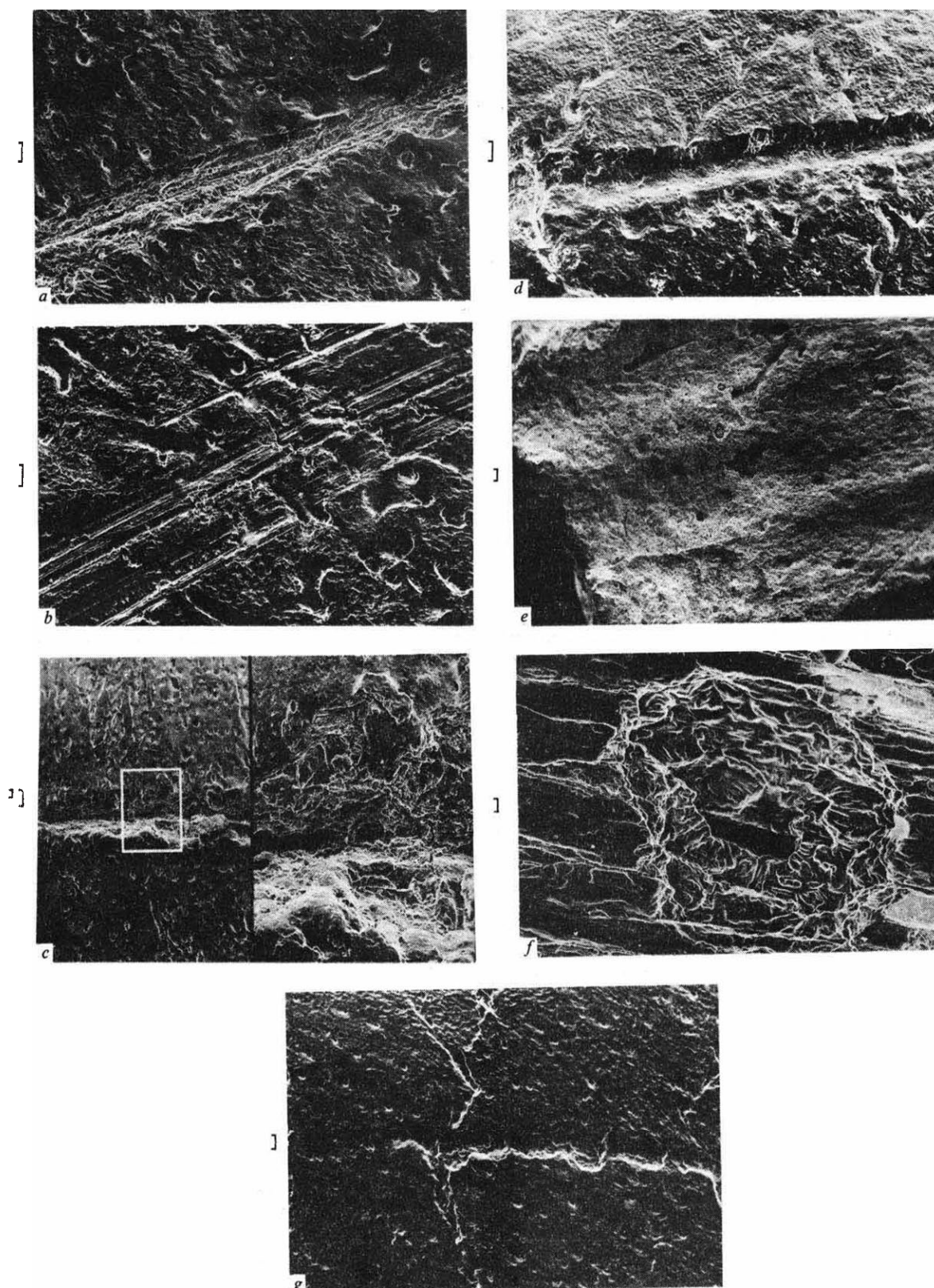


Fig. 1 Micrographs of various tooth and tool marks. The bar or bars to the left of each figure represent 100 μm . *a*, A slicing mark on a modern bone, showing the fine parallel striations within the main groove. *b*, A scraping mark on a modern bone, showing fine striations across a broad area. *c*, A chopping mark on a modern bone, shown at two magnifications. Left, at low power; note the broad, V-shaped groove lacking fine striations. Right, at higher power; note the fragments of bone crushed inwards at the bottom of the groove. *d*, A carnivore tooth scratch on a modern bone. Although this is the functional equivalent of a slicing mark, there are no fine striations. *e*, A rodent gnawing mark on a modern bone. This is the functional equivalent of a scraping mark, but gnawing produces only broad, shallow grooves without fine striations. *f*, A puncture produced by a carnivore canine on a modern bone. Punctures are functionally equivalent to chopping marks and are similar in that fragments of bone are crushed inwards. However, punctures are typically rounded in outline, unlike chopping marks. *g*, A preparator's marks on a fossil bone, produced by an action similar to slicing. The lack of striations distinguish it from a slicing mark and the irregular edges of the groove distinguish it from a tooth scratch.

Table 1 Data on frequencies and locations of different types of marks on fossils from Beds I and II, Olduvai Gorge

Type of mark	No. of specimen	Locations	Taxa	Sites
Slicing	13	Metatarsal shaft (2), radius shaft (2), limb shaft fragment (2), humerus shaft, metacarpal shaft, proximal ulna, distal metacarpal, rib shaft, tibial crest, proximal radius	<i>Kobus sigmoidalis</i> , <i>Equus oldowayensis</i> , giraffid indeterminate, small bovid indeterminate, large bovid indeterminate, medium/large mammal indeterminate	DK Levels 2 and 3, FLK NN Levels 2 and 3, FLK 'Zinj', MNK Main
Scraping	5	Scapula blade, humerus shaft, radius shaft, metatarsal shaft, limb shaft fragment	<i>Kobus sigmoidalis</i> , <i>Equus oldowayensis</i> , giraffid indeterminate, medium mammal indeterminate, large mammal indeterminate	DK Level 3, FLK 'Zinj', MNK Main
Chopping	5	Metacarpal shaft (2), metapodial shaft, proximal phalanx shaft, limb shaft fragment	<i>Parmularius altidens</i> , <i>Equus oldowayensis</i> , bovid indeterminate, mammal indeterminate	DK Levels 2 and 3, FLK NN Level 2, MNK Main
Tooth/gnawing	48	Limb shaft fragment (12), proximal rib (6), humerus shaft (4), rib shaft (3), femur shaft (3), tibial crest (2), metacarpal shaft (2), proximal metacarpal (2), scapula glenoid, proximal ulna, radius shaft, radio-ulna shaft, mandibular corpus, hyoid, parietal vault, proximal femur, acetabulum, distal tibia, fibula shaft, metatarsal shaft, cuneiform, indeterminate fragment	<i>Kobus sigmoidalis</i> , <i>Megalotragus kattwinkeli</i> , <i>Elephas recki</i> , <i>Deinotherium</i> sp.; <i>Equus oldowayensis</i> , <i>Cercocebus</i> , <i>Homo habilis</i> , Tragelaphini indeterminate, Reduncini intermediate, Antilopini indeterminate, medium/large bovid indeterminate, medium/large mammal indeterminate, large mammal indeterminate, very large mammal indeterminate, mammal indeterminate	DK Levels 2 and 3, FLK NN Level 2 and 3, FLK 'Zinj', FLK N Level 6, FLK N Deino. Level, MNK Main, TK Upper Floor
Preparation/excavation	7	Scapula blade, cervical vertebra, rib shaft, astragalus, cuneiform limb shaft fragment, indeterminate fragment	<i>Elephas recki</i> , Alcelaphini indeterminate, medium bovid indeterminate, medium/large Carnivora indeterminate, mammal indeterminate	FLK 'Zinj', FLK NN Level 6
Undetermined	7	Rib shaft, distal tibia, patella, metatarsal shaft, lateral cuneiform, limb shaft fragment, indeterminate fragment	<i>Equus oldowayensis</i> , small bovid indeterminate, very small mammal indeterminate, large mammal indeterminate	DK Level 3, FLK NN Level 2, FLK N Level 6, MNK Main

Seventy-five surfaces were examined; 10 of these possessed marks which classified into two different types. Numbers in parentheses after skeletal location signify the number of times different specimens showed a particular type of mark in that location.

Another likely explanation is that hominids were the primary agent responsible for accumulating and damaging the bones at FLK NN, Level 2, but did not discard stone tools there.

Cutmarks were commonly found on fragments of fossilized bone, particularly on pieces of limb bone shaft, which suggests that the marks occurred during defleshing or dismembering of carcasses with stone tools (Table 1). However, the cutmarks on bones identifiable to skeletal element and species are neither frequent enough nor do they have sufficient regularity of location to suggest a consistent butchery pattern such as can be reconstructed for various species' carcasses much later in time in the New World^{12,13}. To determine the number of cutmarks in a single taxon, all marks found on all postcrania ($N = 138$) of one species, *Equus oldowayensis*, were examined; only 2.9% showed cutmarks. If one or more standard strategies for butchery of equids with stone tools existed, that sample is too small to reveal them.

Specimens from identified skeletal elements were divided into meat-bearing and non-meat-bearing bones. Table 2 shows a preferential occurrence of tooth marks on meat-bearing bones; cutmarks occur about equally on both kinds of bone. Evidently, hominid use of stone tools on bones was not directed solely

towards meat acquisition. Carnivores and rodents, however, focused more directly on meat or on tissues inside meat-bearing bones.

Our evidence supports Leakey's conclusion¹ that there is a direct causal association between the stone artefacts and fossilized bones at many Olduvai sites. This evidence is important because it comes from some of the oldest known excavated archaeological levels: for example, DK I, Level 3, is older than 1.79 ± 0.03 Myr (ref. 2). Furthermore, the cutmarks indicate that hominids processed carcasses to obtain animal material such as meat, ligament or bone. However, this evidence does not bear directly on questions about the frequency of meat-eating and hunting among early hominids¹⁴⁻¹⁷. Moreover, the occurrence of both cutmarks and tooth marks on the bones documents an overprinting of taphonomic events; both hominids and other animals modified the bones in these assemblages. Therefore we cannot attribute the observed patterns of faunal and skeletal representation solely to hominid activity.

This study establishes: (1) a technique for distinguishing cutmarks made with stone artefacts from similar marks not involving early hominid activities; (2) that cutmarks were produced by hominids using stone artefacts on animal tissues at these sites at least as early as 1.79 Myr ago; (3) that these assemblages were exposed to multiple agents of damage and accumulation and thus their composition cannot be taken as indicative solely of hominid behaviours; (4) that not all bones altered by hominids will be found in association with stone tools; and (5) that hominids and carnivores were in competition for carcasses or bones, perhaps to obtain different substances.

We thank the governments of Kenya and Tanzania, Mary and Richard Leakey, and the National Museums of Kenya for permission to study fossils and bones in their care; and Alan Walker and Glynn Isaac for valuable comments. We also thank

Table 2 Occurrence of stone tool cutmarks and animal tooth marks on meat-bearing bones (major limb bones and axial elements) as opposed to non-meat-bearing bones (metapodials, podials, phalanges)

	Stone cutmark	Tooth mark
Meat-bearing bone	10	27
Non-meat-bearing bone	9	7

The sample consists of 53 fossil surfaces from Olduvai, identified to skeletal element. $\chi^2 = 4.15$, degrees of freedom = 1, $0.025 < P < 0.05$.

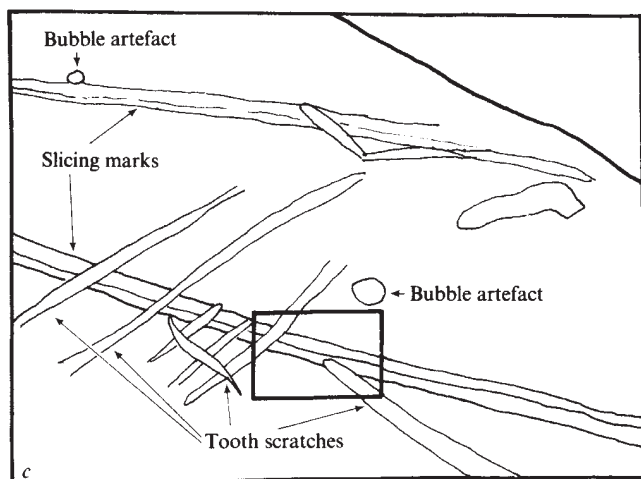
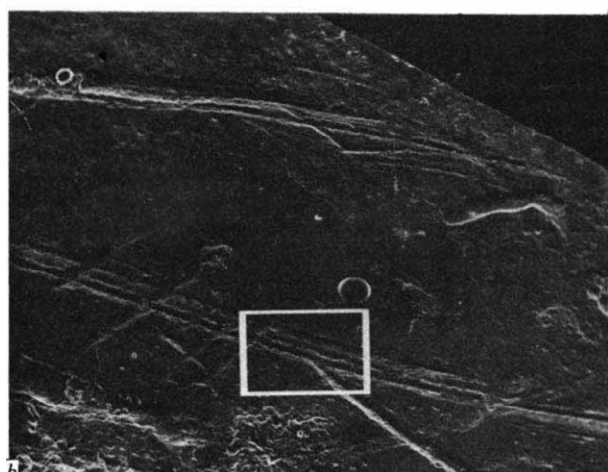
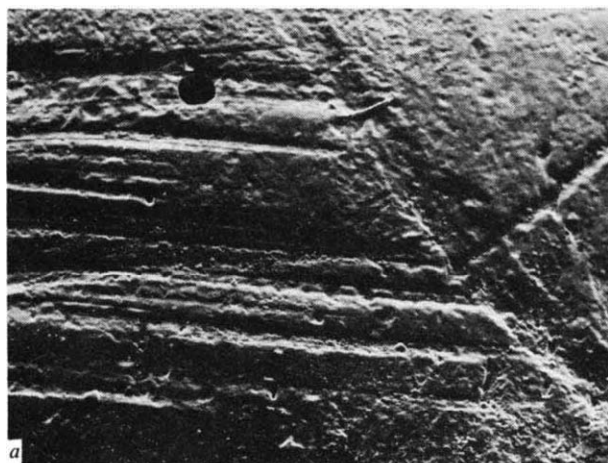


Fig. 2 Marks on Olduvai fossils. *a*, A series of slicing and scraping marks run roughly horizontally across the micrograph. A slicing mark made later in time runs diagonally downwards across the earlier marks. The specimen is an indeterminate shaft fragment from DK I, Level 3. *b*, Two slicing marks run roughly horizontally across this micrograph. The lower one is overlaid by several tooth scratches which must have been made after the slicing marks. The box encloses the intersection of two of these tooth scratches with the slicing mark. The specimen is an equid tibia from FLK Zinj. *c*, A tracing of the micrograph in *b*, with various features labelled.

Linda Perez, Wendy Bosler and Karen Lee Davis for technical assistance. This work was funded by NSF grant BNS 78 19174 to R.P., the Boise Fund (to R.P. and P.S.) and a NIH biomedical research grant 5 S07RR07041-13 to P.S.

Received 18 December 1980; accepted 9 April 1981.

1. Leakey, M. D. *Olduvai Gorge* Vol. 3 (Cambridge University Press, London and New York, 1971).
2. Hay, R. L. *Geology of the Olduvai Gorge* (University of California Press, Berkeley, Los Angeles and London, 1976).
3. Shipman, P. *Ann. N.Y. Acad. Sci.* (in the press).
4. Lanyon, L. E. & Smith, R. N. *Acta orthop. scand.* **41**, 238–248 (1970).
5. Evans, F. G. *Mechanical Properties of Bone* (Thomas, Springfield, 1973).
6. Lakes, R. & Saha, S. *Science* **208**, 501–503 (1979).
7. Bonnichen, R. *Pleistocene Bone Technology in the Beringium Refugium Archaeological Surv. Canada* **89** (1978).
8. Sadek-Kooros, H. *Am. Antiq.* **37**, 369–382 (1972).
9. Brain, C. K. in *Human Origins* (eds Isaac, G. L. & McCown, E.) 97–116 (Benjamin, Menlo Park, 1976).
10. Walker, A. *Phil. Trans. R. Soc. B* **292**, 56–64 (1981).
11. Potts, R. (in preparation).
12. Guilday, J. E., Parmelee, P. W. & Tanner, D. P. *Penna Arch.* **32**, 59 (1962).
13. Frison, G. C. *Plains Anthropol.* **16**, 258 (1971); *The Casper Site* (Academic, New York, 1974).
14. Isaac, G. L. *Scient. Am.* **238** (4), 90–108 (1978).
15. Tanner, N. & Zihlman, A. *Signs* **3**, 585–608 (1976).
16. Binford, L. R. *Nunamiut Ethnoarchaeology* (Academic, New York, 1978).
17. Jolly, C. J. *Man* **5**, 5–20 (1970).

Chondroitin sulphate from fossilized antlers

J. E. Scott & E. W. Hughes

Department of Medical Biochemistry, University of Manchester Medical School, Oxford Road, Manchester M13 9PT, UK

If biopolymers could be isolated from archaeological specimens there would be good prospects of correlating the gene products from different populations of the same species, for example, according to their immunological specificities. Soft tissues are very poorly preserved after only decades, but the macroscopic appearance of bone and antler can remain unchanged for thousands of years. Methods which allow the extraction of relatively undegraded biopolymers from such materials would extend by several orders of magnitude the period from which biopolymers could be identified¹. We report here that chondroitin sulphate, a polysaccharide characteristic of connective tissue and with chemical similarities to immunologically active glycosaminoglycans, has been isolated in good yield from fossilized antlers (3,000–130,000 yr old) which were demineralized by a new procedure. There are now prospects for using biochemical and immunological methods *in situ* and on isolated biopolymers to investigate the composition and mobility of populations of species in the past.

It is first necessary to remove the mineral. Previous methods have used acids or aqueous media containing EDTA but neither can be recommended to secure high yields of undegraded biopolymers of the types which offer maximum prospects of serological specificity, that is, glycoproteins. Aqueous EDTA demineralizes slowly, allowing degradation of water-soluble glycosaminoglycans extracted into solution. We developed a method², using EDTA in organic solvents, which leaves most water-soluble polymers *in situ*. Such media are much less reactive than aqueous environments. Several glycosaminoglycan components of bone and antler were identified³ which had escaped detection in aqueous EDTA-treated samples. Structural proteins and glycoproteins were demonstrable by immunofluorescent techniques in bone and antler demineralized by the new method (J. E. S., R. Timpl and K. von der Mark, unpublished results). It seemed that the application of this method to fossils might give a maximum yield of minimally degraded material.

Table 1 Analyses per g of original antler

Sample	Age (yr)	Ca (nmol)	Hydroxyproline (mg)	Uronic acid in isolated polysaccharide (μ g)	Hexosamine in isolated polysaccharide (μ g)
Modern		5.4*	42	300	
a	3,000–5,000	8.4*	1.65	28.9	38
b	3,800		1.3	21	18.7
c	38,850†	6.7*	17.4	43.4	45‡
d	12,000	6.1§	14.8	110	85
e	130,000	8.2§	0.77	7.8	4.4

* EDTA titration.

† Radiocarbon date (supplied by Mr R. Burleigh).

‡ Galactosamine, by ion exchange chromatography (from Dr A. Shuttleworth).

§ Atomic absorption (supplied by Dr G. A. Lumb).

Two of the samples of antler used were; *a*, red deer (*Cervus elaphus*) from Grimes Graves, estimated as 3,000–5,000 yr old, and *c*, reindeer (*Rangifer tarandus*) from Creswell Crags, Derbyshire, radiocarbon-dated as 38,850 $\pm$ 2,500 yr old. The probable age of sample *b* (red deer antler, find no. 806 from Grimes Graves, Greenwells Pit, in contact with chalk rubble) was 3,800 yr (ref. 4); that of samples *d* (*Megaceros giganteus*, from Antrim bog) and *e* (*Cervus elaphus*) were ~12,000 and 130,000 yr, respectively. These samples were decalcified for 4–8 weeks in 80–20% ethanol–water solutions containing 0.5 M trimethylammonium EDTA at pH 8 and 37 °C (ref. 1). The decalcified material was digested with papain at 65 °C and acidic polysaccharides were isolated from the digest using cetylpyridinium as described elsewhere⁵, and electrophoresed³. Hydroxyproline, hexosamine and uronic acid determinations, and chondroitinase digestion of the isolated polysaccharides were done using methods described elsewhere^{6–9}. Table 1 lists the results.

Electrophoresis, chondroitinase digestion, hexosamine and uronic acid determinations indicated a considerable yield of very good quality chondroitin sulphate, which agrees well with the finding of chondroitin-4-sulphate in contemporary samples of antler³. The yield from antler *c* was ~15% of that from modern antler, on a weight basis.

The success of the methods applied to the fossilized antlers suggests that much older material might yield significant quantities of glycosaminoglycan-type material. Although chondroitin sulphate represents side chains rather than the whole proteoglycan, it would be remarkable if every vestige of the polypeptide had been detached in the fossil. Indeed, amino acids (including serine) typical of the polypeptide core around the attachment region, which are normally found in papain-digested chondroitin proteoglycan, were present in the hydrolysate of chondroitin sulphate from sample *c*. It is possible that some of the chondroitin sulphate constituted part of macromolecules which could be defined as proteoglycan.

The conditions in which the sample was lying should influence the yield of glycosaminoglycan. Sample *c*, from a dry and slightly alkaline limestone cave, was well preserved. More interesting is the excellent yield from sample *d*, which was considerably less mineralized than the others. It is possible that components of the bog water had had a preservative or fixative action.

We observed the histological distribution of chondroitin sulphate in decalcified sample *c* which had been embedded in wax and cut on conventional microtomes, using standard techniques (Alcian Blue–critical electrolyte concentration (CEC) staining according to the method of Scott and Dorling¹⁰). It should be possible to apply immunofluorescence methods to decalcified fossils, as we have to modern antler (J. E. S., R. Timpl and K. von der Mark, unpublished results). Other calcified tissues should be expected to respond, thus making the

'molecular museums' of skeletal tissues available to modern immunological procedures.

Antibodies have been applied to fossils to identify proteins (particularly collagen and albumin¹¹) and it has proved possible to distinguish between fossil species. Glycoproteins carry antigenic reactivities of very high specificity (for example, blood groups) which reside in the glycosaminoglycan portion. The demonstration that a glycosaminoglycuronan structure survives in good yield for tens of thousands of years suggests that it will be rewarding to look for similar compounds, either by immunological or newly developed chemical methods of the required specificity and sensitivity (see part 2 of ref. 2, 6–201). Intraspecies information on populations from the distant past should then be available.

We thank Mr Burleigh for sample *b* and for radiocarbon dating, Professor Martin Jope for samples *d* and *e*, and Dr Prag for samples *a* and *c*; all provided valuable discussion. We also thank Dr A. Shuttleworth for the amino acid determination.

Received 29 September 1980; accepted 8 April 1981.

1. Jope, E. M. *Biogeochemistry of Aminoacids* (ed. Hare, P. E.) 23 (Wiley, New York, 1980).
2. Scott, J. E. & Kyffin, T. W. *Biochem. J.* **169**, 697–701 (1978).
3. Scott, J. E., Kyffin, T. W. & Hughes, E. *Proc. 5th. Int. Symp. Glycoconjugates* (eds Schauer, R. et al.) 566 (Georg Thieme, Stuttgart, 1979).
4. Burleigh, R., Hewson, A., Meeks, N., Sieveking, G. & Longworth, I. *Radiocarbon* **21**, 41–47 (1979).
5. Scott, J. E. *Meth. biochem. Analysis* **8**, 145–197 (1960).
6. Woessner, J. F. Jr *Archs Biochem. Biophys.* **93**, 440–447 (1961).
7. Bitter, T. & Muir, H. M. *Analyt. Biochem.* **4**, 330–334 (1962).
8. Kraan, J. G. & Muir, H. M. *Biochem. J.* **66**, 55p (1957).
9. Saito, H., Yamagata, T. & Suzuki, S. *J. biol. Chem.* **243**, 1536–1542 (1968).
10. Scott, J. E. & Dorling, J. *Histochemie* **5**, 221–233 (1965).
11. Lowenstein, J. M. *Phil. Trans. R. Soc.* (in the press).

Monoclonal antibodies that distinguish between New World species of *Leishmania*

Diane McMahon Pratt & John R. David

Department of Medicine, Harvard Medical School, and Department of Rheumatology and Immunology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA

The New World *Leishmania* species¹ *Leishmania braziliensis*, which causes mucocutaneous leishmaniasis, can be distinguished from *Leishmania mexicana*, which causes cutaneous leishmaniasis, by the buoyant density of nuclear and kinetoplast DNA², electrophoretic mobility of different isoenzymes^{3–6}, by differing characteristic growth patterns in the sandfly and hamster and *in vitro*^{7,8}, and by size in electron microscopy^{9,10}. However, there is no available method for rapid and accurate diagnosis of fresh isolates of New World leishmaniasis species. Furthermore, immunological cross-reaction of *Leishmania* with *Trypanosoma cruzi*, which is co-endemic with the former, has posed a serious problem in the use of serodiagnostic tests for leishmaniasis^{11–13}. The ability to distinguish at primary isolation between the species causing cutaneous and mucocutaneous leishmaniasis would make it possible to concentrate on treating those patients likely to suffer serious disfiguring disease. To develop such a specific diagnostic test, we have produced monoclonal antibodies¹⁴ specific for either the *L. mexicana* or *L. braziliensis* species. One of these monoclonal antibodies is specific for a subgroup of *L. braziliensis*; none is cross-reactive with *T. cruzi* (Y strain) epimastigotes. These monoclonal antibodies should be useful in the taxonomic identification of different species of New World leishmaniae as well as for the direct diagnosis of leishmaniasis.

Female BALB/c mice were injected with membrane-rich preparations isolated from either *L. braziliensis panamanensis* or *L. mexicana amazonensis* promastigotes. The membrane preparations were obtained by disrupting the promastigotes by

nitrogen cavitation, followed by subfractionation by differential centrifugation and purification on sucrose density gradients. The method of fusion was that of Kennett *et al.*¹⁵. Antibody production was assessed between days 14 and 17 after fusion by an indirect radioactive binding assay using leishmanial membranes hydrophobically fixed to PVC microtitre plates. Cultures secreting antibody were then assessed for cross-reactivity with *Leishmania hertigi* or *Leishmania enrietti* antigens. Selected wells were then cloned by limiting dilution.

The average values for six fusions were: 73% of the total wells plated showed hybrid growth; 18% of these hybrid-containing wells produced antibodies to *Leishmania* antigens and 0.6% of the hybrid-containing wells were producing species-specific antibody. The specificity of the monoclonal *Leishmania* antibodies produced is shown in Fig. 1. Monoclonal antibodies IV-3B9, V-1A12 and IX-1G3 are cross-reactive positive controls. The amount of radioiodinated rabbit anti-mouse immunoglobulin that bound to plates containing the *L. braziliensis panamanensis* membrane antigens, previously incubated with the *L. braziliensis* species-specific hybrid supernatants, was 17 to 21 times greater than that which bound after preincubation with control medium. After incubation with *L. mexicana* species-specific hybrid supernatants, the amount of radioiodinated rabbit anti-mouse immunoglobulin that bound to plates containing *L. mexicana amazonensis* membrane antigens was 6.1 to 38 times greater than that which bound to the medium controls. When *L. hertigi* membrane antigens were used, the amount of radioiodinated rabbit anti-mouse immunoglobulin bound in the case of either *L. braziliensis* or *L. mexicana*

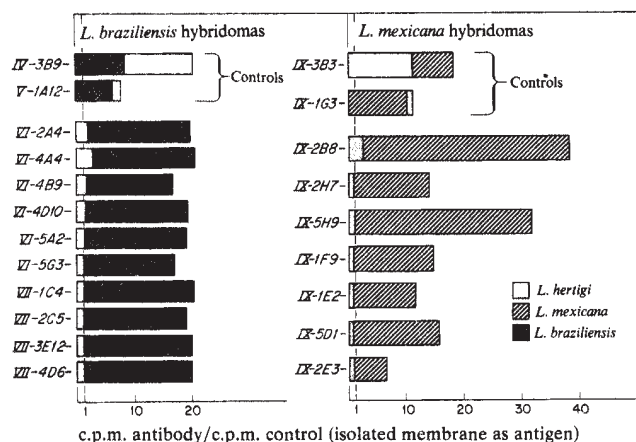


Fig. 1 An indirect radioimmune binding assay was used to determine the reactivities of the culture supernatants from *L. braziliensis*-specific hybrids with *L. braziliensis panamanensis* and *L. hertigi* membrane antigens and the reactivities of culture supernatants from *L. mexicana*-specific hybrids with *L. mexicana amazonensis* and *L. hertigi* membrane antigens. Antibodies produced by the hybridomas were measured by radioactive binding assay. Briefly, the screening assay was as follows. Sonicated leishmanial membranes diluted in phosphate-buffered saline (PBS) containing 0.02% Na₂S₂O₃ were plated in Cooke 'U'-bottom 96-well PVC plates overnight at 5 °C. The plates were washed six times in PBS containing 5% horse serum at 0.02% Na₂S₂O₃. Culture medium supernatants containing secreted antibodies were incubated with the antigen plates for a minimum of 2 h at 5 °C. Immune mouse serum (diluted 1:20) and cell culture medium were used as positive and negative controls, respectively. The plates were again washed six times with PBS containing 5% horse serum and 0.02% Na₂S₂O₃. Affinity-purified ¹²⁵I-labelled rabbit F(ab')₂ anti-mouse immunoglobulin (5–10 ng per well; 10⁵ c.p.m.) was added to the wells and incubated for 1 h at 0 °C. Excess antibody was again removed by washing five times. The plates were then air dried and the radioactivity bound to each well measured using a Packard Auto-Gamma counter. The background (c.p.m. bound) was 510 for *L. braziliensis panamanensis* and 217 for *L. hertigi*, in the case of the *L. braziliensis* hybridomas. The background (c.p.m. bound) was 340 for *L. mexicana amazonensis* and 440 *L. hertigi* for the *L. mexicana* hybridomas.

Table 1 Relative reactivities of *Leishmania* monoclonal antibodies with epimastigotes of *T. cruzi*

Code	Clone	Ratio of bound antibody to bound control		
		<i>L. mexicana amazonensis</i>	<i>L. braziliensis panamanensis</i>	<i>T. cruzi</i>
M-1	IX-2B8-D2	21.0	—	1.1
M-2	IX-2H7-E10	23.0	—	0.7
M-3	IX-5H9-C1	17.0	—	0.9
M-4	IX-1F9-A9	10.6	—	0.9
M-5	IX-1E2-D12	8.7	—	1.3
M-6	IX-2E3-C10	2.4	—	1.0
M-7	IX-5D1-C10	1.4	—	1.1
B-1	VI-4A4-A8	1.0	40.0	1.4
B-2	VI-4B9-D10	1.2	59.7	1.0
B-3	VI-4D10-D12	1.5	39.9	1.3
B-4	VI-2A5-A4	1.4	42.0	1.2
B-5	VII-2C5-C5	1.2	49.0	0.8
B-6	VII-1C4-H6	1.4	45.0	0.9
B-7	VI-2A4-E3	1.2	40.7	1.7
B-8	VII-3E12-D3	1.2	43.0	0.9
B-9	VII-4F11-A6	1.6	38.8	1.0
B-10	VII-4D6-E8	1.2	62.9	0.8
B-11	VI-5G3-F3	1.1	50.0	0.9
	V-1A12-G2	21.5	14.8	14.2
	V-3G5-C6	21.5	8.2	12.4

Reactivities of the culture supernatants of cloned hybrids with external membrane components of *Leishmania* were determined in an indirect radioactive binding assay using glutaraldehyde-fixed *Leishmania*. Glutaraldehyde fixation of the various *Leishmania* strains was carried out as described elsewhere¹⁶. The binding assay procedure was essentially as described in Fig. 1 legend except that 'V'-bottom microtitre plates (Cooke) were used and the *Leishmania* (3–5 × 10⁶ per well) were washed only three times after being pelleted at 400g for 10 min after each incubation with antibody. On completion of the assay, the *Leishmania* were resuspended after the final wash in 100 µl of 0.5 M NaOH and 60 µl was taken for radioactivity measurement.

species-specific supernatant was only 0.7 to 3.0 (average value 1.3) times that bound in the medium controls. Similar specificity was seen if the *L. braziliensis* hybrids were assayed for reactivity with *L. enrietti*, *L. mexicana amazonensis* or *L. hertigi*, or if the *L. mexicana* hybrids were assayed for reactivity with *L. braziliensis panamanensis* or *L. hertigi* (data not shown).

The reactivities of the culture supernatants of cloned hybrids with external membrane components of *Leishmania* were determined in an indirect binding assay using glutaraldehyde-fixed *Leishmania* (Table 1). All the specific hybridoma antibodies tested, except M-7, bound to the appropriate glutaraldehyde-fixed *Leishmania*. As almost all the antibodies of interest appeared to react with external surface components, glutaraldehyde-fixed organisms were used in the indirect binding assay. In addition, this assay was chosen because it avoids proteolysis, which can occur during membrane isolation. In the case of clone M-7, the antigen recognized is either destroyed by glutaraldehyde or the antigen is an internal membrane-associated component.

No significant reactivity was found for any of the species-specific monoclonal antibodies with *T. cruzi* (strain Y) epimastigotes (Table 1). Clones V-1A12-G2 and V-3G5-C6 produce cross-reactive *Leishmania* monoclonal antibodies that are clearly cross-reactive with the *T. cruzi* epimastigotes. The levels of reactivity of these clones with the *T. cruzi* epimastigotes are comparable to those with *L. mexicana amazonensis* and *L. braziliensis panamanensis*.

The species-specific monoclonal antibodies were assessed for their reactivity against a panel of *L. mexicana* and *L. braziliensis* strains (Table 2). Each strain designation had been recently determined by isoenzyme analysis. Isolates designated as WR303, Tres Bracos 04 and Tres Bracos 11 are *L. mexicana*; M2903 and Tres Bracos 05, 08 and 13 are *L. braziliensis*.

Table 2 Comparative reactivities of monoclonal antibodies to strains of *L. mexicana* and *L. braziliensis*

Code	Clone	Ratio of bound antibody to bound control						
		<i>L. mexicana</i>			<i>L. braziliensis</i>			
		WR303	Tres Bracos 04	Tres Bracos 11	Lbb M2903	Tres Bracos 05	Tres Bracos 13	Tres Bracos 08
M-1	IX-2B8-D2	19.6	14.7	21.1	0.8	1.0	1.1	0.9
M-2	IX-2H7-E10	21.4	14.8	17.9	0.6	0.9	0.7	0.7
M-3	IX-5H9-C1	16.0	7.2	10.3	0.6	0.9	0.8	0.7
M-4	IX-1F9-A9	9.3	5.5	7.7	0.6	0.9	0.9	0.8
M-5	IX-1E2-D12	8.1	5.1	7.3	0.8	1.2	1.2	1.0
B-1	VI-4A4-A8	1.0	2.8	ND	15.5	ND	ND	ND
B-2	VI-4B9-D10	1.2	0.9	1.2	7.3	29.5	14.0	13.1
B-3	VI-4D10-D12	1.5	1.9	1.5	60.4	30.9	13.1	12.8
B-4	VI-2A5-A4	1.4	1.2	2.6	7.5	20.4	10.7	6.6
B-5	VII-2C5-C5	1.1	0.7	1.3	78.6	55.2	7.4	14.5
B-6	VII-1C4-H6	1.4	1.1	2.1	5.9	30.3	7.6	10.9
B-7	VI-2A4-E3	1.2	1.6	12.8	10.8	30.6	16.4	14.4
B-8	VII-3E12-D3	1.1	0.9	9.6	63.8	27.5	9.7	11.8
B-9	VII-4F11-A6	1.6	1.2	13.7	4.3	24.5	10.0	93
B-10	VII-4D6-E8	1.2	0.9	5.2	80.2	29.6	14.4	15.7
B-11	VI-5G3-F3	1.1	0.7	1.7	0.9	1.4	5.4	0.4

The reactivities of the culture supernatants of cloned hybrids with glutaraldehyde-fixed isolates of *L. mexicana* and *L. braziliensis* were determined in an indirect radioactive binding assay as described in Table 1 legend. Walter Reed strain WR303 is *L. mexicana amazonensis* isolate LV72 from Liverpool; Tres Bracos strains 04, 05, 08, 11 and 13 are Universidade de Brasilia stocks LTB-0016 (Irene), LTB0042 (Jurandir), LTB0014 (Mucosa), LTB0055 (Jose) and LTB0048 (Elaine) respectively. Strain Lbb M2903 is the reference *L. braziliensis braziliensis* isolate of Dr R. Lainson. Isoenzyme analyses of each strain were performed by Dr M. Miles. ND, not determined. Results of note are demarcated by boxes.

Clones M-1 to M-5 are reactive with the *L. mexicana* strains but unreactive with the *L. braziliensis* strains. In contrast, clones B-1 to B-6 and B-11 are reactive with the *L. braziliensis* strains, but not *L. mexicana* strains. Clones B-7 to B-10 are reactive with all the *L. braziliensis* strains but not the *L. mexicana* strains *L. mexicana amazonensis* WR303 and Tres Bracos 04; however, these are reactive with the *L. mexicana* strain Tres Bracos 11. Clone B-11 seems to be specific for a subgroup of *L. braziliensis*, reacting only with the strains *L. braziliensis* (Tres Bracos 13) and *L. braziliensis panamanensis*.

The cross-reactivity of clones B-8, B-9 and B-10 may be useful in studying the subspeciation of *L. mexicana* strains. It is interesting that the *Leishmania* strains Tres Bracos 04 and Tres Bracos 11 were isolated from the same area of Brazil (Tres Bracos). In addition, clone VI-5G3-F3, reacting with *L. braziliensis* Tres Bracos 13, may be useful in studies of subspeciation of *L. braziliensis*. Work is being done to develop these monoclonal antibodies as clinical reagents for direct diagnosis of leishmaniasis.

We thank Elizabeth Bennett for technical assistance and Dr Larry Hendricks for advice and for providing the *L. mexicana amazonensis* (WR303) and *L. braziliensis panamanensis* (WR120) strains. We also thank Dr R. Lainson for the *L. braziliensis braziliensis* reference strain M2903, Dr Michael Miles for isoenzyme analysis, Dr Rodney Hoff for *T. cruzi* epimastigotes and Dr Philip Marsden for the Tres Bracos strains 04, 05, 08, 11 and 13 and for discussions which directly led to the work. This work was supported by NIH grant AI-16479 and by grants from the Rockefeller foundation and CNPQ, the National Research Council of Brazil.

Received 22 December 1980; accepted 25 March 1981.

- Marsden, P. *New Engl. J. Med.* **300**, 350–352 (1979).
- Chance, M. L., Peters, W. & Shchory, L. *Ann. trop. Med. Parasit.* **68**, 307–316 (1974).
- Gardener, P. J., Chance, M. L. & Peters, W. *Ann. trop. Med. Parasit.* **68**, 317–325 (1974).
- Kilgour, V., Gardener, P. J., Godfrey, D. G. & Peters, W. *Ann. trop. Med. Parasit.* **68**, 245–246 (1974).
- Martin, E., Simon, M. W., Schaefer, F. W. III & Muhhada, A. J. *J. Protozoology* **23**, 600–607 (1978).
- Miles, M. A., Pova, M. M., de Souza, A. A., Lainson, R. & Shaw, J. J. *Trans. R. Soc. trop. Med. Hyg.* **74**, 243–252 (1979).
- Lainson, R. & Shaw, J. J. *Br. med. Bull.* **28**, 44–48 (1972).

- Zuckerman, A. & Lainson, R. in *Parasitic Protozoa* Vol. 1 (ed. Kreier, J. P.) 57 (Academic, New York, 1977).
- Shaw, J. J. & Lainson, R. *J. trop. Med. Hyg.* **79**, 9–13 (1976).
- Gardener, P. J., Shchory, L., & Chance, M. L. *Ann. trop. Med. Parasit.* **71**, 147–155 (1977).
- Camargo, M. E., Hoshino, S., Correa, N. S. & Peres, B. A. *Rev. Inst. Med. Trop. Sao Paulo* **13**, 45–50 (1971).
- Neal, R. A. & Miles, R. A. *Rev. Inst. Med. Trop. Sao Paulo* **12**, 325–332 (1970).
- Camargo, M. E. & Robonato, C. *Am. J. trop. Med. Hyg.* **18**, 500–505 (1969).
- Köhler, G. & Milstein, C. *Nature* **256**, 495 (1975).
- Kennett, R. H., Denis, K. A., Tung, A. S. & Klinman, N. R. in *Curr. Topics Microbiol. Immun.* **81**, 77 (1978).
- Williams, A. F. *Eur. J. Immun.* **3**, 628–632 (1973).

Induction of a light requirement during seed development and its ecological consequences

E. G. Cresswell & J. P. Grime

Unit of Comparative Plant Ecology (NERC), Department of Botany, The University, Sheffield S10 2TN, UK

Germination in many arable weeds and herbaceous plants of marshland and heathland depends on exposure of the seed to unfiltered sunlight^{1–4}. In contrast, seeds of many grasses, legumes and woody species exhibit no such requirement and are capable of germination in darkness^{1–4}. These interspecific differences are apparent in seeds immediately after their removal from the parent plant and in theory could be due to major differences in seed physiology. Here we present evidence supporting the alternative hypothesis that the contrasting responses to light are imposed by differences in the light-filtering properties of the maternal tissues which surround the developing seeds.

Induction of a light requirement has been observed when seeds develop on plants grown under incandescent lighting⁵ and when mature detached seeds receive light filtered through a leaf canopy^{6–14}. These effects are the result of a decrease in the red to far-red ratio of the light incident on the seed and are mediated by the phytochrome system^{15–17}. As seeds of most species are

formed initially within green photosynthetic tissues, which only later senesce and dehisce, it is possible that, in natural conditions, a similar process of induction may take place while the developing seed is still attached to the inflorescence.

Photoconversion of phytochrome occurs only when the seed is moist¹⁸. A critical factor is, therefore, whether the chlorophyll content of the structures investing the seed is retained throughout maturation, or is lost before the seed is dry. In the former case the seeds will receive only filtered light with a low red/far-red ratio, whereas in the latter the tissues will lose the capacity to act as a filter and the red/far-red ratio of the light experienced by the developing seed will increase. As the seed matures its phytochrome will be arrested in a particular photo-stationary state, determined by the light environment present immediately before drying out. This state controls the light response of the freshly dispersed seed. Thus, seeds which have matured entirely within green tissues would have most of their phytochrome in the inactive (Pr) form, requiring a light stimulus for germination. In contrast, seeds which have been exposed to unfiltered light before drying regardless of the light quality received at earlier stages of development, would have their phytochrome in the active (Pfr) form and be able to germinate in light or in darkness.

To investigate this hypothesis, we have studied the ripening process and germination requirements of seeds removed from natural populations of species commonly occurring in Britain. In each species, seed heads providing a range of developmental stages, from newly formed to mature seed, were collected and divided into five age classes. Determination of the moisture content of the seeds in each class provided an index of maturity, and the filtering capacity of the investing structures at each stage was estimated by extraction and measurement of chlorophyll.

The investing structures were then removed in daylight from a subsample of the seeds of each maturity stage and germination tests were performed in light (16 h day, 40.0 W m⁻² fluorescent supplemented by tungsten lighting, red/far-red 7.6) and darkness in a growth room¹⁹ with 20°C day and 15°C night temperatures. The remaining intact seeds were then dried in darkness over CaCl₂ for ~6 weeks, after which the investing structures were removed and the germination tests repeated.

The results suggest that species conform to one of two broad types. The first, represented in Fig. 1 by *Atriplex hastata*, *Stachys*

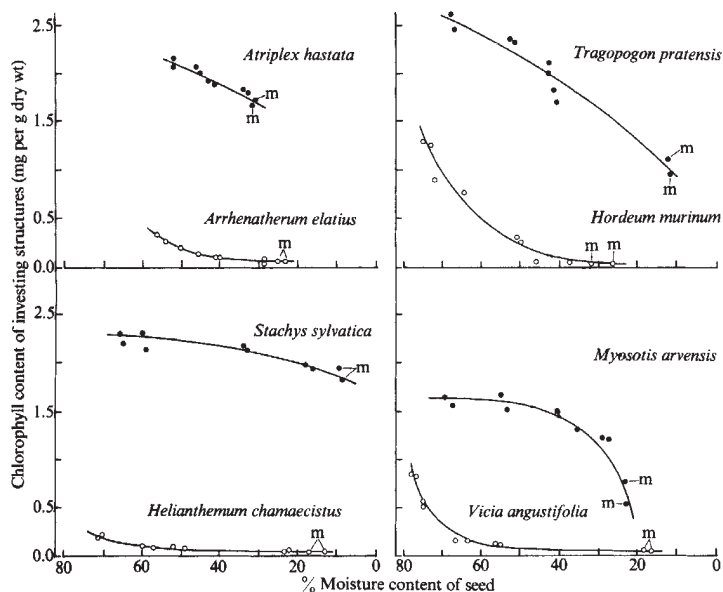


Fig. 1 Relationship between the moisture content of seeds and the chlorophyll concentration of the investing structures at various stages of inflorescence development in eight species of flowering plants. ●, Species producing mature seeds with a light requirement⁴; ○, species producing mature seeds with no light requirement⁴; m, mature seeds.

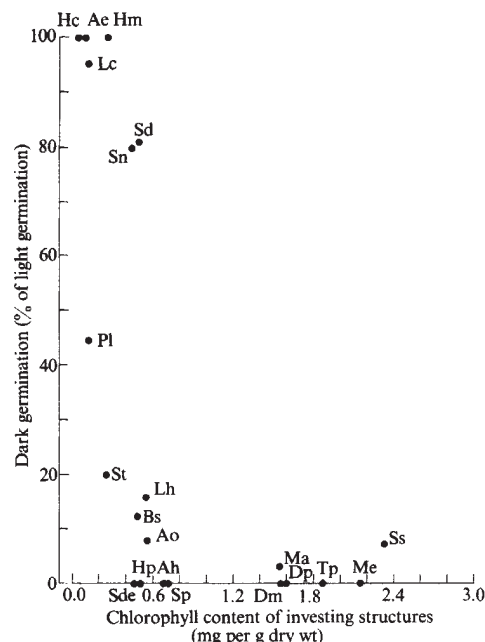


Fig. 2 Relationship between chlorophyll concentration of the investing structures and dark germination of mature seeds in various species of flowering plants. Values for chlorophyll refer to the concentration at the mid-point in the range of seed moisture contents. Ae, *Arrhenatherum elatius*; Ah, *Arabidopsis hirsuta*; Ao, *Anthoxanthum odoratum*; Bs, *Brachypodium sylvaticum*; Dm, *Draba muralis*; Dp, *Digitalis purpurea*; Hc, *Helianthemum chamaecistus*; Hm, *Hordeum murinum*; Hp, *Hypericum perforatum*; Lc, *Lotus corniculatus*; Lh, *Leontodon hispidus*; Ma, *Myosotis arvensis*; Me, *Milium effusum*; Pl, *Plantago lanceolata*; Sd, *Silene dioica*; Sde, *Sieglingia decumbens*; Sn, *Silene nutans*; Sp, *Succisa pratensis*; Ss, *Senecio squalidus*; St, *Serratula tinctoria*; Tp, *Tragopogon pratensis*.

sylvatica, *Tragopogon pratensis* and *Myosotis arvensis*, includes species in which chlorophyll is retained in the investing structures throughout seed maturation. The second is composed of species such as *Arrhenatherum elatius*, *Hordeum murinum*, *Helianthemum chamaecistus* and *Vicia angustifolia* in which chlorophyll declines to a low level early in seed development.

When the germination of mature dry-stored seeds of some of these species is examined (Fig. 2) it is evident that the retention of chlorophyll in the investing structures is strongly correlated with the presence of a light requirement.

Figure 3 shows that the light response may change during maturation and can be manipulated experimentally. *Hordeum murinum* is a species which will germinate to a comparable extent in light and darkness at any stage of maturity if the seed coats are removed in full daylight before drying the seed and conducting the germination tests. If, however, the seed is dried in darkness before removal of the investing structures, seed of the first two maturity stages will not germinate in darkness. This reveals the presence in immature seeds of *H. murinum* of an hitherto undetected phase in which a light requirement is imposed by the green parental tissues. During the normal course of maturation this requirement is lost, because the chlorophyll content declines sharply (Fig. 1) when the seed is still moist enough for its phytochrome system to respond to the resulting changes in the quality of the light transmitted to the seed. Similarly, Fig. 3 shows that the obligate light requirement of *Tragopogon pratensis* may be removed if unripe seeds are taken from their green investing bracts while still moist, and exposed to unfiltered light before undergoing the dark test. However, if these seeds are dried in darkness with the investing structures intact, the light requirement is maintained, as in the normal process of maturation within the green bracts.

These different types of response may have important ecological implications. Plants which retain chlorophyll in the investing structures throughout the maturation process tend to

be the small-seeded species known to develop more or less persistent reserves of seeds buried in the soil²⁰⁻²⁵. The light requirement conferred on these seeds by their maturation environment may enable them to remain dormant when buried. Rather different are the larger-seeded composites, such as *Tragopogon pratensis*, *Leontodon hispidus* and *Senecio squalidus*, all of which retain chlorophyll in their bracts and exhibit strong dark inhibition of germination. Here, there is no evidence that the seeds become buried or form persistent reserves. One possible explanation, now under investigation, is that seeds released with phytochrome in the inactive form may be unusually susceptible to canopy-induced dormancy. This might allow the exploitation of vegetation gaps by delaying germination until unfiltered daylight is perceived by seeds lying on the soil surface. A similar sensitivity to canopy-induced dormancy may contribute to the delaying mechanism which allows small seeds to become buried and incorporated into persistent reserves rather than germinating soon after seed fall. An extreme example of chlorophyll retention is found in some of the ruderal species, such as *Chenopodium album* and *Atriplex hastata*, where the investing structures also maintain a high moisture content and ripe seeds occur within green fleshy tissues. In addition to its effect on the light requirement of the seed, this phenomenon may attract predation and thus facilitate seed dispersal.

Species in which chlorophyll declines at an early stage include two conspicuous groups: the large-seeded, autumn-germinating grasses and the legumes. In both, water rather than light seems to provide the environmental cue for germination. Autumn-germinating grasses, for example *Arrhenatherum elatius*, *Festuca arundinacea* and *Helictotrichon pratense*, which produce and release their seeds over the summer months, germinate more or less synchronously with the onset of wet weather in the autumn, regardless of their light environment²⁵⁻²⁷. The legumes, on the other hand, are held dormant by their impervious seed coats and will germinate in light or darkness once this barrier to imbibition is overcome^{4,28,29}.

From this study we predict that the light requirement of seeds may be strongly affected by (1) seed maturity at the time of

collection, (2) removal or retention of investing structures, and (3) the light conditions during post-harvest drying. Differences in harvesting and drying procedures may account for much of the conflicting data in the literature concerning the light requirements of seeds, and differential rates of chlorophyll loss and/or seed drying within the same inflorescence may explain the germination 'polymorphism' with respect to light requirement reported for certain species³⁰.

We thank S. R. Band for his interest and assistance in the investigation. This research was supported by the Natural Environment Research Council.

Received 27 November 1980; accepted 13 March 1981.

1. Kinzel, W. *Frost und Licht als beeinflussende Kräfte der Samenkeimung* (Ulmer, Stuttgart, 1920).
2. Toole, V. K. *Seed Sci. Technol.* **1**, 339-396 (1973).
3. Grime, J. P. & Jarvis, B. C. in *Light as an Ecological Factor II* (eds Bainbridge, R., Evans, G. C. & Rackham, O.) 525-532 (Blackwell, Oxford, 1975).
4. Grime, J. P. *et al. J. Ecol.* **69** (in the press).
5. McCullough, J. M. & Shropshire, W. *Pl. Cell Physiol.* **11**, 139-148 (1970).
6. Borthwick, H. A., Hendricks, S. B., Toole, M. W. & Toole, V. K. *Bot. Gaz.* **115**, 205-225 (1954).
7. Black, M. in *Dormancy and Survival* (ed. Woolhouse, H. W.) 193-217 (Cambridge University Press, 1969).
8. Pigott, C. D. *New Phytol.* **70**, 953-966 (1971).
9. Smith, H. in *Seed Ecology* (ed. Heydecker, W.) 219-231 (London, 1973).
10. Gorski, T. *Physiol. Pl.* **34**, 342-346 (1975).
11. King, T. L. *New Phytol.* **75**, 87-90 (1975).
12. Silvertown, J. W. *New Phytol.* **85**, 109-118 (1980).
13. Fenner, M. *New Phytol.* **84**, 95-102 (1980).
14. Fenner, M. *New Phytol.* **84**, 103-106 (1980).
15. Flint, L. H. & McAlister, E. D. *Smithsonian misc. Collns* **94**, 1-11 (1935).
16. Flint, L. H. & McAlister, E. D. *Smithsonian misc. Collns* **96**(2), 1-8 (1937).
17. Borthwick, H. A., Hendricks, S. B., Parker, M. & Toole, V. K. *Proc. natn. Acad. Sci. U.S.A.* **38**, 662-666 (1952).
18. Evenari, M. *Handb. Pfl. Physiol.* **15**, 804-847 (1965).
19. Rorison, I. H. *New Phytol.* **63**, 358-362 (1964).
20. Brenchley, W. E. *J. agric. Sci.* **9**, 1-31 (1918).
21. Chippindale, H. G. & Milton, W. E. *J. Ecol.* **22**, 508-531 (1934).
22. Milton, W. E. *J. Ecol.* **27**, 149-159 (1939).
23. Champness, S. S. & Morris, K. J. *Ecol.* **36**, 149-173 (1948).
24. Wesson, G. & Wareing, P. F. *J. exp. Bot.* **20**, 401-413 (1969).
25. Thompson, K. & Grime, J. P. *J. Ecol.* **67**, 893-921 (1979).
26. Chancellor, R. J. *Proc. 9th Br. Weed Control Conf.*, 1129-1135 (1968).
27. Davison, A. W. *J. Ecol.* **59**, 493-506 (1971).
28. Quinliven, B. J. *Aust. J. agric. Res.* **12**, 1009-1022 (1961).
29. Ballard, L. A. T. *Seed Sci. Technol.* **1**, 285-303 (1973).
30. Cavers, P. B. & Harper, J. L. *J. Ecol.* **54**, 307-382 (1966).

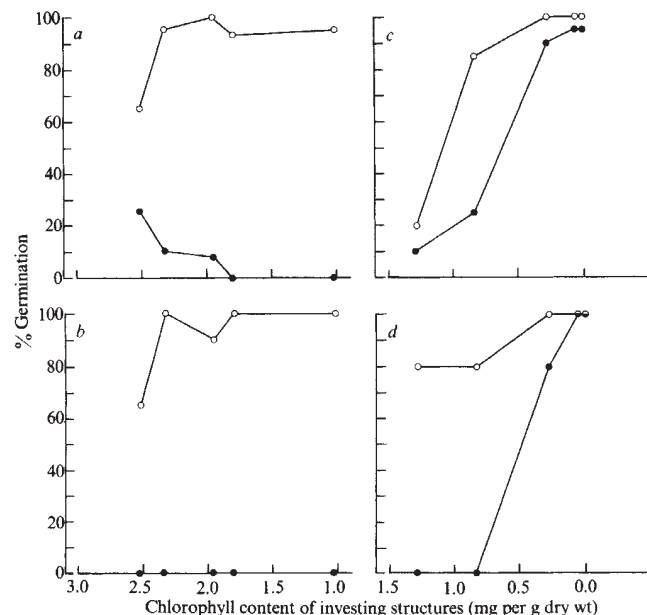


Fig. 3 Germination in light (○) and darkness (●) of seeds of *Tragopogon pratensis* (a, b) and *Hordeum murinum* (c, d). In both species, seed maturity is characterized by reference to the chlorophyll content of the investing structures. Dry storage involved placing the intact inflorescence over calcium chloride in darkness for ~6 weeks. All germination tests on fresh (upper panel) and dry-stored (lower panel) seeds were preceded by removal of the investing structures, a procedure involving ~5-min exposure to daylight.

Depression of net photosynthesis in *Vicia faba* L. exposed to sulphur dioxide and ozone

D. P. Ormrod, V. J. Black & M. H. Unsworth

Department of Physiology and Environmental Studies, University of Nottingham, Sutton Bonington, Loughborough LE12 5RD, UK

In most laboratory investigations of plant responses to air pollution, plants are exposed to a single pollutant gas. However, in rural atmospheres, pollutants seldom occur singly; for example, in central England, where summertime mean concentrations of sulphur dioxide are typically 10-30 p.p.b. (parts in 10⁹ by volume) depending on the site, there are 10-30 days each summer when concentrations of ozone generated photochemically exceed 80 p.p.b. for at least 1 h (ref. 1). Maximum hourly ozone concentrations of 130-260 p.p.b. have been reported¹⁻³. Similarly, maximum hourly SO₂ concentrations in rural central England exceed 50 p.p.b. for a few hours each month. We have previously shown⁴ that net photosynthetic rates P_N of field beans, *Vicia faba* L. c.v. Dylan, were depressed, in comparison with rates in clean air, on exposure to SO₂ concentrations as low as 40 p.p.b. Responses of photosynthesis to mixtures of SO₂ and O₃ seem to have been studied only at higher concentrations atypical of the natural environment^{5,6}. We report here that the addition of ozone in the concentration range 60-150 p.p.b. to atmospheres containing 40 p.p.b. SO₂ led to decreases in P_N of *V. faba* L. larger than those observed in either gas alone.

Three-week old plants of *V. faba* L. c.v. Dylan, grown in a controlled environment were exposed for 4 h in chambers described fully elsewhere⁷ at $22 \pm 2^\circ\text{C}$, 280 W m^{-2} photosynthetically active radiation, and $0.9 \pm 0.2 \text{ kPa}$ vapour pressure deficit. Air flowed through the chambers at approximately 20 l min^{-1} and internal fans minimized boundary layer resistance.

Photosynthetic rates of plants exposed to pollutants were compared with those of 'control' plants in an identical chamber containing clean air. Before exposure, plants were acclimatized for 24 h with clean air in both chambers. During the period immediately before exposure to pollutants, rates of photosynthesis were compared between chambers; typical rates were $2 \text{ g CO}_2 \text{ m}^{-2} \text{ h}^{-1}$. Then air containing either 40 p.p.b. SO_2 or O_3 or $\text{O}_3 + 40 \text{ p.p.b. SO}_2$ was supplied to one chamber and P_N was monitored in both chambers. There were 18 experiments with O_3 and 21 with $\text{O}_3 + \text{SO}_2$; in each experiment three plants were exposed in one pot in each chamber.

When either O_3 or $\text{O}_3 + \text{SO}_2$ was added, P_N decreased. The maximum depression of P_N during exposure increased with increasing O_3 concentration. The change δP_{NT} between P_N of treated plants on the day before exposure and the minimum P_N during exposure was assumed to be related linearly to the change δP_{NC} between photosynthetic rates of control plants on the two days, and to O_3 concentration. Least-squares regression analysis was used to fit the expression

$$\delta P_{NT} = a\delta P_{NC} + b[\text{O}_3] + c \quad (1)$$

where a , b and c are constants, to the measurements with O_3 alone and with $\text{O}_3 + \text{SO}_2$. The value of a , which is a measure of the influence of environmentally and genetically determined responses of P_N , was held constant between the two sets of observations. The value of b is the rate of decrease of P_N with O_3 concentration [units: $\text{g CO}_2 \text{ m}^{-2} \text{ h}^{-1} (\text{p.p.b. O}_3)^{-1}$]. The value of c [units: $\text{g CO}_2 \text{ m}^{-2} \text{ h}^{-1}$] gives the depression of P_N at zero O_3 concentration. For the treatment with O_3 alone, c therefore indicates the magnitude of residual variation in P_N not accounted for by variation in the control. For the treatment with $\text{O}_3 + \text{SO}_2$, c also gives the depression in P_N caused by SO_2 alone.

Values of b and c differed between the two treatments and an analysis of variance was used to establish the statistical significance of the differences. Table 1 shows the values of a , b and c for the $\text{O}_3 + \text{SO}_2$ treatment, and the differences from values derived for O_3 alone. The differences are highly significant ($P < 0.01$).

The regression coefficients indicate that photosynthesis was significantly depressed by 40 p.p.b. SO_2 alone (a point reported previously⁴), but the addition of increasing concentrations of O_3 to a background of 40 p.p.b. SO_2 depressed photosynthesis further. The rate of decrease in P_N with increasing O_3 concentration was larger in O_3 alone than in $\text{O}_3 + 40 \text{ p.p.b. SO}_2$.

Figure 1 illustrates the regression equations for the case $\delta P_{NC} = 0$. At low O_3 concentrations, the depression of P_N was greater (δP_{NT} more negative) in $\text{O}_3 + \text{SO}_2$ than in O_3 alone. At

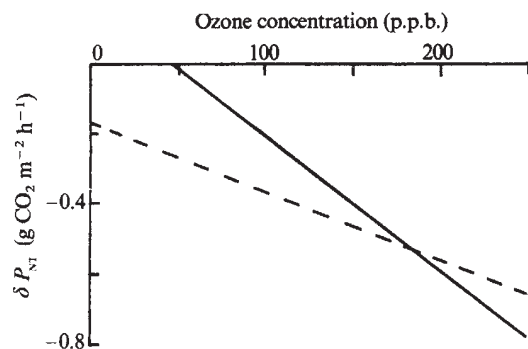


Fig. 1 Regressions of the depression of net photosynthesis δP_{NT} against ozone concentration for *V. faba* plants exposed to O_3 alone (—) or to $\text{O}_3 + 40 \text{ p.p.b. SO}_2$ (---). $\delta P_{NT} = (P_N \text{ during exposure}) - (P_N \text{ before exposure})$.

Table 1 Coefficients a , b and c with standard errors in equation (1)

	$\delta P_{NT} = a\delta P_{NC} + b[\text{O}_3] + c$	
	$\text{O}_3 + 40 \text{ p.p.b. SO}_2$	Difference Δ
a	0.393 ± 0.224	—
b	-0.0019 ± 0.0004	-0.0020 ± 0.0007
c	-0.170 ± 0.068	0.345 ± 0.110

δP_N is the difference in net photosynthetic rate during exposure to pollutants and before exposure. Suffices T and C refer to plants exposed to pollutants and control plants respectively. Values of a , b and c for $\text{O}_3 + \text{SO}_2$ exposures are shown; differences Δ between b and c for the O_3 and $\text{O}_3 + \text{SO}_2$ exposures are also shown, such that $b(\text{O}_3) = b(\text{O}_3 + \text{SO}_2) + \Delta$ and so on. Units: a is non-dimensional; b , $\text{g CO}_2 \text{ m}^{-2} \text{ h}^{-1} (\text{p.p.b. O}_3)^{-1}$; c , $\text{g CO}_2 \text{ m}^{-2} \text{ h}^{-1}$.

ozone concentrations above about 150 p.p.b., the two treatments produced similar depressions in P_N . It may be significant that the threshold for visible injury by O_3 in these experiments was at about this concentration.

These results suggest that photosynthesis which is depressed by SO_2 at concentrations commonly occurring in Britain would be depressed further by the superposition of ozone generated photochemically. This effect would be most important near urban areas with sources of appropriate pollutants and precursors.

On the basis of these results we propose the following general mechanism for responses to mixtures of pollutant gases. At low concentrations of pollutants the injury done by one pollutant may be insufficient to influence an independent response to a second pollutant. High concentrations of one pollutant may lead to substantial injury which masks or over-rides any response to a second pollutant. Interactions between pollutants and subsequent synergistic responses seem most likely to be observed at concentrations between these extremes.

This work was supported by NERC and SRC. We thank H. L. Back for advice on statistical analysis.

Received 30 September 1980; accepted 10 April 1981.

1. Martin, A. & Barber, F. R. *Atmos. Envir.* (in the press).
2. Bell, J. N. B. & Cox, R. A. *Envir. Pollut.* **8**, 163–170 (1975).
3. Apling, A. F. *et al. Nature* **269**, 569–573 (1977).
4. Black, V. J. & Unsworth, M. H. *J. exp. Bot.* **30**, 473–480 (1979).
5. Bennett, J. H. & Hill, A. C. in *Air Pollution Effects on Plant Growth* (ed. Dugger, M.) ACS Symp. Ser. **3**, 115–127 (American Chemical Society Washington, 1974).
6. Carlson, R. W. *Envir. Pollut.* **18**, 159–170 (1979).
7. Black, V. J. & Unsworth, M. H. *J. exp. Bot.* **30**, 81–88 (1979).

Gene transfer in *Nicotiana rustica* using irradiated pollen

J. L. Jinks, P. D. S. Caligari* & N. R. Ingram

Department of Genetics, University of Birmingham, Birmingham B15 2TT, UK

In spite of reports of widespread matromorphy—the occurrence of diploid plants whose genomes are exclusively maternal in origin—in *Nicotiana*¹ and in other genera^{2,3}, attempts to induce it in pure-breeding material of *Nicotiana rustica* have met with little success^{4,5}. When heavily irradiated pollen from one variety is used as the inducer on a contrasting variety acting as maternal parent, the progeny, while showing a greater resemblance to the maternal parent, are not true matromorphs because they usually have at least one characteristic of the paternal parent. At the highest dose of γ -irradiation, 20 krad, the progeny usually

* Present address: Scottish Crop Research Institute, Pentlandsfield, Roslin, Midlothian EH25 9RF, UK.

Table 1 Five characters recorded for the maternal (V27) and paternal (V12) parents, the F₂ of the cross between these, and the corresponding M₂ obtained by self-pollination of M₁ plants

	Plant colour Yellow/green	Inflorescence Mop/non-mop	Height at flowering time (cm)	Height 10 weeks after planting (cm)	Final height (cm)
V27	25:0	25:0	77.02±1.18	110.96±1.02	120.60±1.10
V12	0:25	0:25	126.46±2.64	163.84±1.60	170.60±2.30
F ₂	2:98	3:97	121.39±2.08	173.68±1.94	180.97±2.30
M ₂	16:81	23:74	93.77±2.85	135.60±3.27	148.27±3.43
Selection of M ₂ plants					
Yellow plants (16/97)	—	10:6	85.06±6.07	115.25±8.23	128.63±8.31
Yellow plants + mop head (10/16)	—	—	79.00±5.65	103.60±9.58	121.70±11.89
Yellow plants + mop head with black pigment (6/10)	—	—	81.17±6.78	114.17±12.65	131.17±12.91

The M₁ plants resulted from pollination of V27 with pollen of V12 that had been γ -irradiated (20 krad). Selections were made from the M₂ population on the basis of recessive yellow plant colour and mop-head inflorescence (typical of V27), and the dominant black-pigmented ovary (typical of V12). Height values are mean±s.e.m.

acquire from the paternal parent only one or a small sample of the many heritable characteristics by which the two varieties differ. By selecting within these progenies we have shown here that lines can be isolated with the characteristics of the pure-breeding maternal variety but with the exception of a specific characteristic transferred from the paternal variety. To achieve the same end by conventional means would require a combination of many generations of recurrent backcrossing and selection.

We have worked with the varieties V12 and V27 of *N. rustica*. For several genetic characteristics (for example, colour) V12 is dominant over V27 and was therefore chosen as the paternal parent. Irradiated pollen (10–20 krad) of V12 was used to pollinate the pure-breeding V27 to yield progeny M₁, from which progeny M₂ were obtained by random self-pollination. We have compared the characteristics of M₂ with those of the progeny F₂ obtained by crossing the two varieties using unirradiated pollen. The M₂ progenies resembled V27 more than did F₂ progenies from the same cross made with unirradiated pollen⁵, the similarity to V27 increasing with the three doses used. However, although all the M₂ families produced from the highest dose closely resembled the initial maternal parent, V27, there was significant phenotypic heterogeneity among them. We concluded that the most plausible explanation was the effective transfer of incomplete, random parts of the genome of the irradiated parent into the M₁ and M₂ progenies whose genomes were, therefore, largely derived from V27, the maternal parent of the M₁ crosses. Further segregational data and cytological examinations are required to test this explanation but none of the usual signs of gross chromosomal abnormalities are present in the M₂.

Because the M₂ families resulting from the highest dose of radiation showed not only the greatest resemblance to V27 but also significant heterogeneity in their phenotypes, there are obviously opportunities for selection among them. It would, for example, be of particular practical significance if we could demonstrate the presence of individuals identical to V27 except for a single characteristic which they have acquired from the pollen parent, V12: that is, if we could demonstrate the practicability of transferring single desirable characters from one genotype into another without the simultaneous acquisition of other less desirable characters which would then have to be bred out by a combination of recurrent backcrossing and selection.

Table 1 gives the results of five characters recorded for the M₂ progenies derived from a cross involving V27 (maternal parent) and pollen of V12 which had received 20 krad of γ rays. Two of these characters, plant colour and inflorescence morphology, display the effects of major gene differences, thus the results are given as segregation ratios. The other three characters, height at flowering time, height 10 weeks after planting and final height, are continuously variable and are therefore expressed as means (cm)±s.e.m. The corresponding results for V27, V12 and their F₂ are also given. For every character the F₂ is more like V12

than V27, that is, V12 has mainly the dominant alleles for these characters.

If irradiation of the V12 pollen had no effect, the results for the M₂ and F₂ generations would differ only by sampling variation, but this was not the case. For each character, the M₂ shows a greater resemblance to the recessive maternal parent, V27. However, the continuously varying characters are correlated, thus the information they provide is not completely independent⁵.

Plant colour is the first character that can be scored during development thus it provides the first criterion for selecting maternal phenotypes. Of the 97 M₂ plants, 16 showed the recessive yellow colour of V27. During the development of these 16 plants, the other characters were scored (see Table 1). Among these plants a high proportion (10 out of 16) also expressed the other recessive major gene-controlled character of V27, mop-head inflorescence. This proportion is surprisingly high as there were no plants with the combination of yellow colour and mop-head inflorescence in the equivalent F₂.

M₂ plants selected for yellow plant colour also had mean heights that were closer to those of the maternal parent, V27, than those of the unselected M₂. If from among the yellow plants of the M₂, those which are also mop heads are selected, they have means which are even closer to those of V27 for height at flowering time and final height. However, the mean height 10 weeks after planting is below even the low mean of V27 because of two particularly low scoring plants in the small sample of 10.

As our intention was to examine the possibility of transferring a paternal character from V12 into the maternal genotype of V27, we next selected individuals from among those 10 plants which expressed the appropriate paternal characteristic: the remaining major gene-controlled difference, namely, the black anthocyanin pigmentation present in the ovary of V12 and absent in V27. Of the 10 plants, 6 had this pigmentation and were therefore of the desired type. Thus from the original 97 M₂ plants it was possible to select six with the maternal phenotype for two major gene controlled characters and paternal phenotype for the third. No plant with this phenotype was present in a sample of 100 F₂ of the same cross.

The means of the quantitative characters (see Table 1) for these six selected plants, although very close to those of the maternal parent, V27, do show a slight deviation when compared with the 10 plants from which they were chosen. This suggests that some factor, although very minor, has been transferred with the pigment gene or alternatively that the latter has a pleiotropic effect on the quantitative characters.

In an earlier experiment (ref. 4 and D. S. Virk and G. S. Pooni, unpublished data) in which only plants selected for yellow colour, mop-head inflorescence and black anthocyanin pigmentation of the ovary were taken beyond the M₁ generation, the presence versus the absence of the pigment segregated in over half of the progenies produced by self-pollinating pigmented plants in the M₂, M₃ and M₄ generations. The progenies

produced by self-pollinating the remaining pigmented plants and all the recessive non-black segregants in the M_2 bred true in all subsequent generations. Therefore, we can readily extract families from the M_2 which are pure-breeding for the recessive maternal traits, yellow plant colour and mop-head inflorescence and the dominant paternal black ovary, although M_3 progenies have to be raised to distinguish between plants which are homozygous, and therefore pure-breeding for black ovary, from those which are heterozygous and therefore still segregating for the recessive non-black ovary.

Because of the greater clarity of the results, we have concentrated in these selection studies on the major gene controlled characters. The same questions, however, can be posed for the quantitative differences between V27 and V12. For example, can plants be selected which are identical to V27 except in respect of one quantitative character for which they are more like V12? Of the 97 M_2 plants, 4 were identical to V27 for all the major gene controlled characters. One of these had a final height of 158 cm, which is more like that of V12, while the other quantitative characters were more like V27. Similarly in the earlier experiment (ref. 4 and D. S. Virk and G. S. Pooni, unpublished data) some of the 120 families which had all the major gene-controlled characteristics of V27 were more like V12 for one of the seven quantitative characters recorded.

It is premature to offer explanations for the similarities and differences between our results and those of Pandey^{6,7}. At this early stage it is more important to establish a working model that allows the routine transfer of one or more characteristics of a paternal parent to progeny that are overwhelmingly like the maternal parent and at the same time will allow the more fundamental questions about mechanisms to be investigated. By selecting in the M_2 in our model system we can transfer specific paternal characteristics into a maternal background with a speed and reliability that would be difficult to match by conventional breeding. This work has already prepared the way for similar transfers between more distantly related paternal and maternal parents which do not produce successful sexual hybrids.

We acknowledge the financial support of the Agricultural Research Council and a postgraduate research studentship from the SRC to N.R.I. We thank Professor Sir Kenneth Mather for his advice and encouragement and Dr D. S. Virk and G. S. Pooni for allowing us to refer to their unpublished data.

Received 10 January; accepted 26 March 1981.

1. Goodspeed, T. H. *The Genus Nicotiana* (Chronica Botanica, Massachusetts, 1954).
2. Eenink, A. H. *Euphytica* **23**, 429–433 (1974).
3. Eenink, A. H. *Euphytica* **23**, 435–445 (1974).
4. Virk, D. S., Dhahi, S. J. & Brumpton, R. J. *Heredity* **39**, 287–295 (1977).
5. Caligari, P. D. S., Ingram, N. R. & Jinks, J. L. *Heredity* (in the press).
6. Pandey, K. K. *Nature* **256**, 310–313 (1975).
7. Pandey, K. K. *Genetics* **49**, 56–69 (1978).

Visual cortex of an anthropoid ape

François Vital-Durand* & Colin Blakemore†

*Laboratoire de Neuropsychologie Expérimentale, INSERM, Unité 94, 16 Avenue du Doyen Lépine, 69500 Bron, France

†University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, UK

Marg et al.¹ recorded activity from an unspecified region of the human occipital cortex but the properties of neurones there were very different from those in cat or monkey striate cortex. Work on man's closest relatives, the apes², might give a better indication of the way in which visual information is analysed in the human cortex. We had an unusual opportunity to study an old, non-reproductive, female black gibbon (*Hylobates concolor*, a lesser Asian ape from the forests of south-east Asia). The properties of neurones in striate and pre-striate cortex in this animal suggest that the visual areas of anthropoid apes are similar in their organization to those of monkeys^{3–5}.

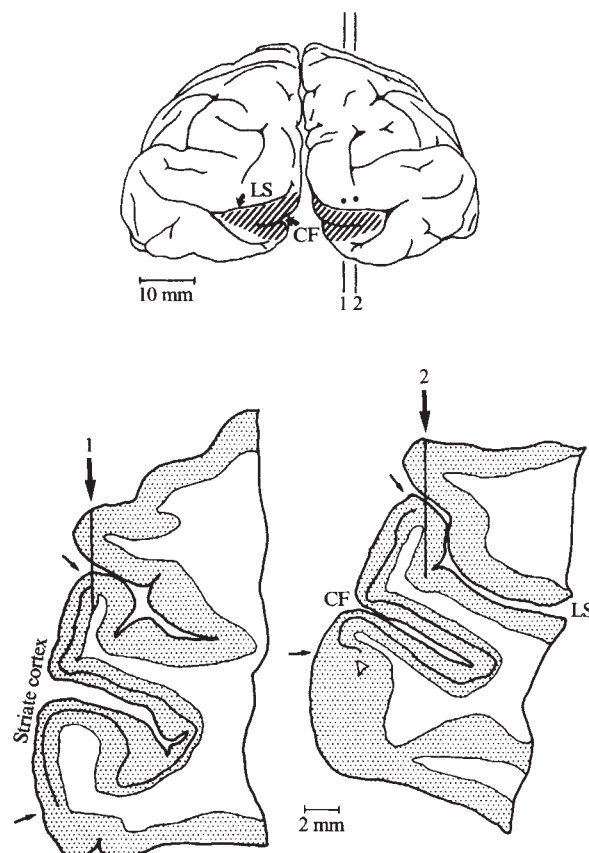


Fig. 1 The top of the figure is a view from behind of the cerebral hemispheres of the gibbon. The exposed part of the striate cortex (area 17) around the calcarine fissure (CF) is shown striped. Filled circles represent the points of entry of the two vertical penetrations, just anterior to the lunate sulcus (LS) in the right hemisphere. The diagrams at the bottom are of sections in the parasagittal planes marked 1 and 2 on the drawing of the brain. The grey matter is stippled and area 17, whose anterior and posterior boundaries are marked with small arrows, is distinguished by the stria of Gennari, shown as a line in the grey matter. Both electrode tracks passed through pre-lunate cortex and across the lunate sulcus: track 1 then re-entered grey matter just anterior to the 17/18 border and ran into the deeper layers of striate cortex; track 2 went down the posterior bank of the lunate sulcus.

The gibbon was premedicated with corticosteroid, tranquilized with ketamine and prepared for an electrophysiological experiment under intravenous (i.v.) steroid anaesthetic (Althesin, Glaxo). During recording, anaesthesia and paralysis were maintained with a continuous i.v. infusion of Nembutal and Flaxedil in glucose-saline solution. The animal was artificially ventilated with air, maintaining end-expiratory CO_2 at $\sim 4.5\%$. The general experimental techniques were similar to those used previously in monkeys⁶. The eyes, protected with contact lenses, were brought to focus, by means of additional spectacle lenses, on a tangent screen at a distance of 114 cm, where we back-projected patterns to investigate the receptive fields of cortical neurones.

The gross disposition of the striate area in the occipital cortex of the gibbon is similar to that of the chimpanzee⁷ where visually evoked potentials have been recorded (C. N. Woolsey and C. Welt, personal communication). Much of the primary visual cortex is buried in the calcarine fissure and only a small strip of striate cortex is exposed on the posterior pole of the hemispheres, around the lips of the fissure (Fig. 1).

We made two, long stereotaxically-vertical penetrations with a glass-insulated tungsten microelectrode⁸ through the occipital cortex of the right hemisphere, starting in pre-striate cortex,

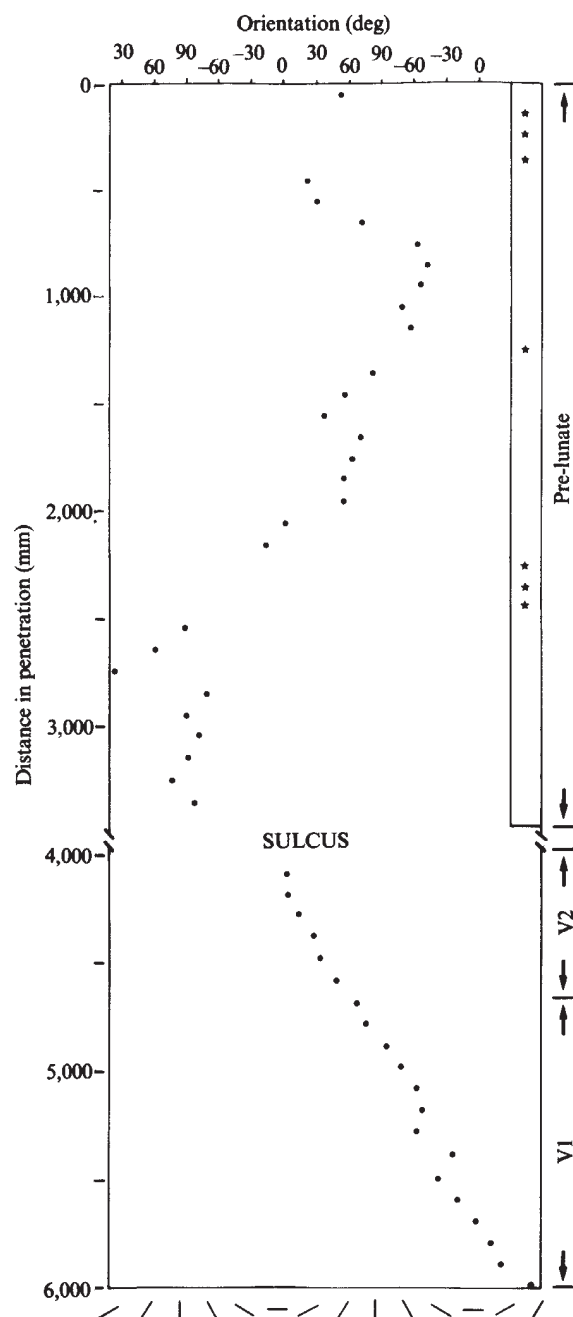


Fig. 2 Properties of neurones recorded in track 1 (see Fig. 1) are reconstructed as a function of distance along the penetration. Orientation-selective cells are plotted as filled circles, the abscissa being the angle of a bar or edge that produced the best response. Non-oriented cells are plotted as stars in the column at the right. The upper part of the track was in pre-lunate cortex, it then crossed the sulcus, entered V2 and passed into V1 (striate cortex). Note the regular sequential shifts in preferred orientation from cell to cell in V1 and V2.

anterior to the lunate sulcus. In both penetrations (see Fig. 1) the electrode passed through this pre-lunate area, across the lunate sulcus and into the anterior lip of the pre-calcarine operculum, just ahead of the cytoarchitectonic border between areas 17 and 18. Track 1 then ran into area 17, while track 2 passed into the posterior bank of the lunate sulcus, which would be occupied by the second visual area in monkeys. All 79 neurones isolated during these penetrations were visually responsive and their receptive fields lay within 4° of the horizontal meridian.

Primary visual cortex (V1): we studied only 14 cells in the striate cortex but nevertheless we were struck by their similarity

to neurones in area 17 of the monkey. All were selective for the orientation of a bar or edge; they could be classified as simple or complex⁹ and some had a preference for short lines. Preferred orientation shifted progressively as the electrode passed obliquely through the cortex (Fig. 2) in the 'columnar' pattern typical of monkeys¹⁰. Receptive fields were small (average size 0.79 deg^2) and they were all close to the vertical meridian, just above the foveal projection. All but one of the neurones were binocularly driven (Fig. 3). Unfortunately the electrode entered the striate cortex just below layer IVc, thus we had no opportunity to determine whether this layer contained non-oriented, monocularly driven cells arranged in ocular dominance clusters: certainly in the chimpanzee¹¹, as in monkeys^{12,13}, the left- and right-eye afferent axons terminate in layer IVc in alternating stripes.

Second visual area (V2): the vertical meridian is represented at the border between area 17 behind, and 18 in front. The retinotopic map is mirror-reversed in the portion of area 18 that runs forward from the border into the posterior bank of the lunate sulcus: the receptive fields of cells encountered in this region, presumably homologous to V2 of the monkey¹⁴, moved from the vertical meridian into the contralateral field as the electrode moved down the posterior bank away from the 17/18 border in track R2. The receptive fields were a little larger than in V1 (mean size 1.41 deg^2) but all but one of the 21 cells isolated in V2 were orientation-selective, there was again evidence of a 'columnar' pattern (Fig. 2) and only one cell was not binocularly driven (Fig. 3).

Pre-lunate visual area: 44 cells were recorded in the pre-lunate area. In both tracks, as the electrode moved down, the receptive fields shifted in towards the vertical meridian from $\sim 11 \text{ deg}$ left to $\sim 6 \text{ deg}$ left. The receptive fields of cells found in the first penetration were very close to the horizontal meridian whereas those for the more lateral track were slightly higher in the visual field. This suggests that there is a topographically organized pre-lunate visual area, for which the representation of the vertical meridian lies somewhere in the anterior bank of the lunate sulcus. This presumably means that this area is not homologous to V3 in the rhesus monkey, as V3 joins V2 at the

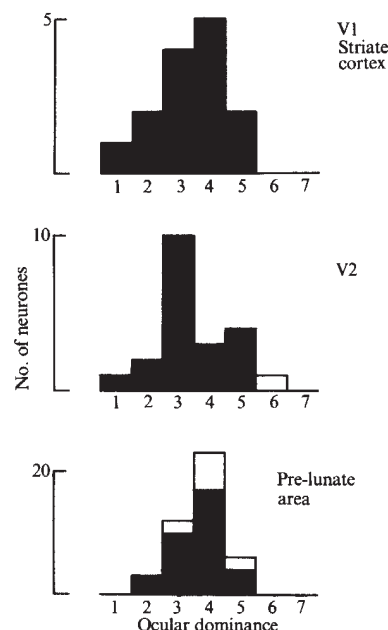


Fig. 3 Ocular dominance histograms⁹ for the three visual areas studied. Cells in groups 1 and 7 are monocularly driven by contralateral and ipsilateral eyes respectively. All other groups represent binocular neurones, with a spectrum of dominance from strong dominance by the contralateral eye (group 2), through equal excitability (group 4), to strong dominance by the ipsilateral eye (group 6). Orientation-selective cells are represented by solid blocks, non-oriented cells by open blocks.

representation of the horizontal meridian¹⁴, but that it may be equivalent to area V3A (ref. 14). This, in turn, implies that there is at least one other visual area between V2 and this pre-lunate area.

The receptive fields of pre-lunate cells were quite large (average size 8.43 deg²), 34 of these (77%) were selective for orientation but the remainder were neither orientation- nor direction-selective. Of the 34 cells with an orientation preference, 18 had some degree of direction selectivity (a slightly higher proportion than in V1 or V2) and each cell encountered was binocularly driven (Fig. 3). Only one cell had a clear colour preference when tested with broad-band coloured stimuli; thus it is unlikely that this area is homologous with the pre-striate V4

complex to which Zeki¹⁵ ascribes the analysis of chromatic information. Indeed we were struck by how few cells in any area had an obvious colour preference.

Our results suggest that areas V1 and V2 in anthropoid apes are organized much as in monkeys and that there is probably a patchwork of further pre-striate areas, stretching forwards towards the parietal and temporal lobes. The close genetic similarity of gibbon to man makes it likely that the human visual areas are similarly arranged.

We thank Dr Vre J. Rinjard for supplying the gibbon, which was being disposed of by the Museum d'Histoire Naturelle de Paris. The work was supported by grants from INSERM and MRC.

Received 28 January; accepted 9 April 1981.

1. Marg, E., Adams, J. E. & Rutkin, B. *Experientia* **24**, 348–350 (1968).
2. Bruce, E. J. & Ayala, F. J. *Nature* **276**, 264–265 (1978).
3. Zeki, S. M. in *Essays on the Nervous System* (eds Bellairs, R. & Gray, E. G.) 327–343 (Clarendon, Oxford, 1974).
4. Allman, J. *Prog. Psychobiol. physiol. Psychol.* **7**, 1–53 (1977).
5. Van Essen, D. C. *A. Rev. Neurosci.* **2**, 227–263 (1979).
6. Blakemore, C., Garey, L. J. & Vital-Durand, F. *J. Physiol., Lond.* **283**, 223–262 (1978).

7. Polyak, S. *The Vertebrate Visual System* (University of Chicago Press, 1957).
8. Merrill, E. G. & Ainsworth, A. *Med. Biol. Engng* **10**, 662–672 (1972).
9. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **195**, 215–243 (1968).
10. Hubel, D. H. & Wiesel, T. N. *J. comp. Neurol.* **158**, 267–293 (1974).
11. Tigges, J. & Tigges, M. *Brain Res.* **166**, 386–390 (1979).
12. Wiesel, T. N., Hubel, D. H. & Lam, D. M. K. *Brain Res.* **79**, 273–279 (1974).
13. Hendrickson, A. E., Wilson, J. R. & Ogren, M. P. *J. comp. Neurol.* **182**, 123–136 (1978).
14. Van Essen, D. C. & Zeki, S. M. *J. Physiol., Lond.* **277**, 193–226 (1978).
15. Zeki, S. *Nature* **284**, 412–418 (1980).

Thrombolysis by human tissue plasminogen activator and urokinase in rabbits with experimental pulmonary embolus

O. Matsuo, D. C. Rijken & D. Collen

Center for Thrombosis and Vascular Research, Department of Medical Research, University of Leuven, Belgium

Thromboembolism is a frequent and often lethal complication of many medical diseases and surgical procedures. Streptokinase, a bacterial non-enzymatic protein, and urokinase, an enzyme obtained from human urine, are both potent activators of the human fibrinolytic system, and have been extensively investigated as a means of medical treatment of this condition. Although it is well established that these agents can clear obstructed vessels in roughly 50% of cases, the fact that they induce relatively extensive systemic fibrinogen breakdown and a serious haemorrhagic diathesis has prevented their general application for thrombolysis. The physiological plasminogen activator in blood (blood activator), which is probably released from the vascular wall (vascular activator) and is identical or at least very similar to the activator extracted from human organs (tissue activator), differs from urokinase (for references see ref. 1). This activator has a markedly higher fibrinolytic to fibrinogenolytic ratio than urokinase *in vitro*¹ and might therefore constitute a better thrombolytic agent. We have developed a purification method for tissue plasminogen activator from a human melanoma cell line which enables us to produce milligramme quantities on a laboratory bench scale² and we have now compared its thrombolytic effect with that of urokinase in an experimental animal model. We have found that tissue plasminogen activator causes thrombolysis at lower doses than urokinase, without extensive plasminogen activation in the circulating blood and without haemostatic breakdown.

The fibrinolytic activities of tissue plasminogen activator and urokinase were determined on plasminogen-enriched bovine fibrin plates³ and expressed in International Units (IU) using the WHO 1st International Reference Preparation of urokinase as described previously⁴. The specific activities of both enzymes were very similar to those described elsewhere². Human fibrinogen was prepared as described by Blombäck and Blombäck⁵ and labelled with ¹²⁵I according to McFarlane⁶.

An experimental thrombus was produced in a polyethylene tube of 3 mm internal diameter by mixing citrated whole human blood (0.5 ml) with 8 µl of ¹²⁵I-labelled human fibrinogen (~10⁶ c.p.m.), 50 µl of 25 mM CaCl₂ and 2 µl of human

thrombin (final concentration 1 NIH unit per ml). After incubation at room temperature for 15 min, 1.5 cm length of the tube (~0.10 ml thrombus) was cut out of the middle position of the tube. The thrombus was poured into a Petri dish, washed repeatedly with 0.15 M NaCl and its ¹²⁵I content measured. The radioactive thrombus was then aspirated in a silicon catheter for injection.

New Zealand White rabbits were anaesthetized with 0.5 ml Hypnorm and 0.3 ml pentobarbital. A jugular vein was dissected free, cut over a distance of 3 mm and, after injection of the thrombus, carefully sutured to restore the blood flow. The activator infusions (tissue plasminogen activator or urokinase) were given intravenously into a contralateral marginal ear vein as a bolus injection of 10% of the total dose followed by a continuous infusion over 6 or 12 h as indicated; blood samples were collected from a femoral vein catheter on citrate at regular

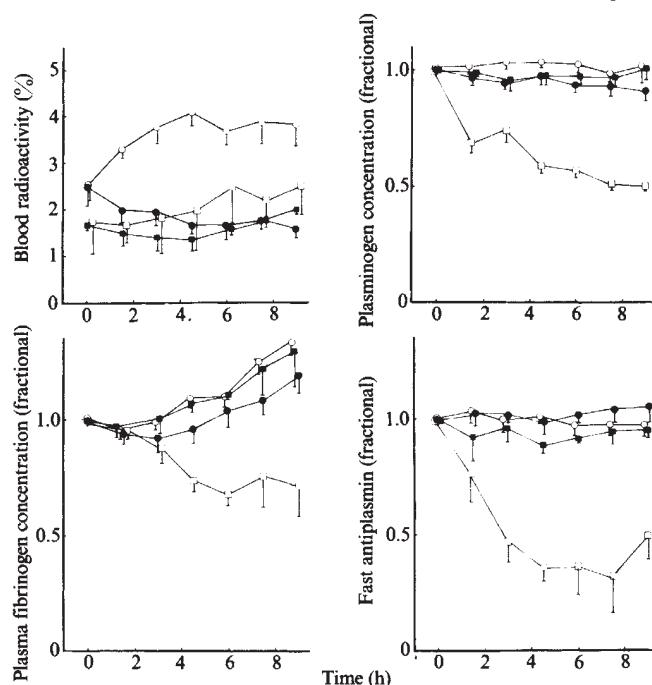


Fig. 1 Changes in the main blood parameters during infusion of tissue plasminogen activator or urokinase in rabbits. ■, Control animals infused with saline; ●, 100,000 units of urokinase in five rabbits; □, 1,000,000 units of urokinase in three rabbits; ○, 35,000 units of tissue plasminogen activator over 6 h in six rabbits and 70,000 units over 12 h in three rabbits. The data represent mean values $\pm$ s.d. Blood radioactivity was expressed as per cent of the injected radioactivity present in the total blood volume.

Table 1 Extent of thrombolysis and isotope recovery

Group	n	% Thrombolysis	% Recovery
Saline 9 h	4	3.2±1.5	102.0±2.4
Saline 24 h	6	5.8±3.2	100.3±4.4
UK 100,000	5	4.0±1.2	103.3±1.1
UK 1,000,000	3	11.5±5.4	97.7±2.9
TA 35,000	6	16.5±2.7	100.4±3.3
TA 70,000	3	23.1±7.4	97.0±0.8

The extent of thrombolysis was determined as the difference in radioactivity between the injected and recovered clots. The isotope recovery was measured as the sum of the radioactivity present in the recovered thrombus, the blood, the urine and the lungs and expressed as per cent of the injected tracer. UK, urokinase; TA, tissue plasminogen activator; n, number of experimental animals. The data represent means $\pm$ s.d.

time intervals for up to 24 h. The animals were killed 9 or 24 h after the start of the infusion, the thorax was opened, the lungs and heart removed together and the thrombus localized with the use of an isotope monitor. The lungs were then carefully dissected, the pulmonary artery incised and the remaining radioactive thrombus recovered. The degree of thrombolysis was calculated as the difference in ^{125}I content of the injected and recovered thrombus. In all but three instances the remaining thrombus was recovered as a single fragment, tightly lodged in a major branch of the pulmonary artery.

Figure 1 represents changes in the main blood parameters during the first 9 h of the infusion. The radioactivity in the blood rose to ~2% of the amount injected in samples taken after 15 min. In the control groups without activator infusion and in the urokinase groups the blood radioactivity remained at this level throughout the experiment, whereas infusion of tissue plasminogen activator (35,000 IU over 6 h or 70,000 IU over 12 h) resulted in a significant rise in blood radioactivity (from 2.2% to a maximum of 4.1%). The fibrinogen level⁵ rose progressively in controls and in those given 10,000 IU urokinase tissue plasminogen activator (infused over 6 h), probably as a result of an acute phase reaction. On infusion of 1,000,000 IU urokinase over 6 h the fibrinogen level fell to 70%. The plasminogen⁷ and fast antiplasmin⁸ levels did not change significantly in the control, 100,000 IU urokinase and tissue plasminogen activator groups, but fell to 50 and 30%, respectively, in the group infused with 1,000,000 IU urokinase.

The extent of thrombolysis and the isotope balance (sum of radioactivity in thrombus, blood, urine and lungs) are represented in Table 1. Significant thrombolysis occurred within 9 h when 1,000,000 IU of urokinase ($P < 0.05$) were infused over 6 h but not with 100,000 IU of urokinase. Infusion of 35,000 IU of tissue plasminogen activator over 6 h resulted in a significant degree of thrombolysis ($P < 0.001$) whereas infusion of 70,000 IU over 12 h yielded significantly more lysis ($P < 0.01$).

The present findings thus confirm and extend previous *in vitro* observations that tissue plasminogen activator has a much higher fibrinolytic to fibrinogenolytic ratio than urokinase and suggest that it might have greater potential as a thrombolytic agent.

Urokinase was a gift from Lepetit (courtesy of Dr Brioude) and from Choay (courtesy of Dr Toulemonde). D. C. R. was recipient of a fellowship from the Netherlands Organisation for the Advancement of Pure Research (Z.W.O.). This study was supported by grants from the Geconcerteerde Onderzoeksacties (Conventie 80/95-3) and from the ASLK Kankerfonds.

Received 21 November 1980; accepted 1 April 1981.

1. Matsuo, O., Rijken, D. C. & Collen, D. *Thromb. Haem.* (in the press).
2. Rijken, D. C. & Collen, D. *J. Biol. Chem.* (in the press).
3. Astrup, T. & Müllertz, S. *Archs Biochem. Biophys.* **40**, 346-351 (1952).
4. Rijken, D. C., Wijngaards, G., Zaaij-De Jong, M. & Welbergen, J. *Biochim. biophys. Acta* **580**, 140-153 (1979).
5. Blombäck, B. & Blombäck, M. *Ark. Kemi* **10**, 415-423 (1956).
6. McFarlane, A. S. *Nature* **182**, 53 (1958).
7. Friberger, P. & Knös, M. in *Chromogenic Peptide Substrates* (eds Scully, M. F. & Kakkar, V. V.) 128-140 (Churchill Livingstone, Edinburgh, 1979).
8. Edy, J., De Cock, F. & Collen, D. *Thromb. Res.* **8**, 513-518 (1976).

Demonstration of a specific dynorphin receptor in guinea pig ileum myenteric plexus

Charles Chavkin & Avram Goldstein

Addiction Research Foundation and Department of Pharmacology, Stanford University, Palo Alto, California 94304, USA

Although the opioid peptide dynorphin contains the pentapeptide Leu-enkephalin as its amino-terminal sequence¹, the two opioid peptides differ in their potency and sensitivity to antagonism by the specific opiate antagonist naloxone. In the guinea pig myenteric plexus-longitudinal muscle assay, dynorphin₍₁₋₁₃₎ is 700 times more potent and 13 times less sensitive to naloxone antagonism than Leu-enkephalin¹. We have identified the specific amino acids in the dynorphin₍₆₋₁₃₎ sequence responsible for this higher potency². One explanation for the difference in naloxone sensitivities is that dynorphin and Leu-enkephalin act through different types of opiate receptors, as first suggested by Goldstein *et al.*¹. This hypothesis has been strongly supported by recent results³ showing that vasa deferentia of mice infused with both D-Ala²-D-Leu⁵-enkephalin (DADLE) and the opiate alkaloid sulfentanil become tolerant to the effects of Leu-enkephalin and sulfentanil, whereas sensitivity to dynorphin₍₁₋₁₃₎ remains unchanged. Here we report that the guinea pig ileum myenteric plexus contains two physically distinct classes of opiate receptors: the μ receptor, with which leu-enkephalin (or normorphine) interacts, and a separate dynorphin receptor.

We used the irreversible antagonist, chlornaltrexamine (CNA), synthesized by Portoghesi *et al.*⁴ to distinguish the different receptors for dynorphin₍₁₋₁₃₎ and for Leu-enkephalin. A specific antagonist should be able to selectively block one receptor population. Alternatively, if CNA were not specific, a specific and reversible ligand should be able to protect one receptor population selectively from inactivation by CNA.

CNA acts as a nonequilibrium opiate receptor antagonist⁵. In agreement with this, we have found that in the guinea pig ileum-longitudinal muscle preparation, the extent of antagonism by CNA depends on both its concentration and the duration of exposure. The effects of 20-min exposure to different CNA concentrations on the dose-response curve of Leu-enkephalin are shown in Fig. 1. Low concentrations of CNA (0.5-3.0 nM) caused parallel shifts of the Leu-enkephalin dose-response curve, a result that is not typical of irreversible inhibition⁷.

To confirm that CNA is irreversible, 10 muscle strips were washed repeatedly after 20-min exposure to 3 nM CNA. After washing for 60 min, the Leu-enkephalin mean inhibitory concentration IC₅₀ was 3.0 μM (16.5-fold increase in IC₅₀), and after an additional 4 h washing, the mean IC₅₀ was 2.4 μM . The Leu-enkephalin IC₅₀ of six muscle strips not exposed to CNA but washed for the same period changed very little (170 nM initially, 270 nM after 5 h). Thus within the time period of the experiment, the antagonism by CNA was irreversible.

Parallel shift of the dose-response curve by an irreversible antagonist indicates that the myenteric plexus contains spare opiate receptors⁷ although there are other possible explanations⁸. Creese and Snyder⁹ concluded that there are no spare receptors in the myenteric plexus because receptor binding affinity (apparent K_d) was well correlated with pharmacological potency (IC₅₀) for a series of non-peptide opiates. Our result might be reconciled with theirs if the parallel shift in the Leu-enkephalin dose-response curve demonstrated here were not due to spare μ receptors (isomorphic spare receptors) but rather due to the presence of receptors of different types (allomorphic spare receptors) in the myenteric plexus.

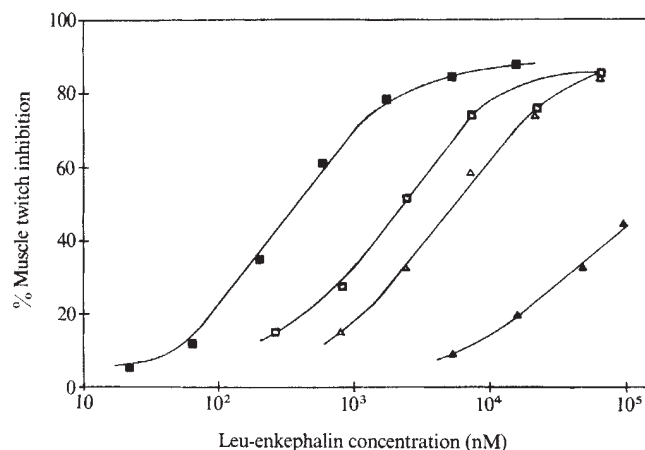


Fig. 1 The effects of different CNA treatments on Leu-enkephalin dose-response curve: ■, before treatment; □, after 0.5 nM CNA; △, after 3 nM CNA; and ▲, after 25 nM CNA. Each point represents the mean per cent inhibition of muscle twitch from 3–4 experiments. Ileum longitudinal muscle from adult male Hartley guinea pigs (Simonsen) was used as described elsewhere⁶. Dilutions of Leu-enkephalin (Biosearch) made in methanol/0.1 M HCL (1:1, v/v) were added to the 5-ml organ bath containing the electrically stimulated muscle strips. A stock solution of 10 μ M CNA (given by P. S. Portoghesi and A. E. Takemori) in ethanol/0.1 M HCL (1:1, v/v) was stored at -70°C and fresh dilutions prepared each day. After 20-min CNA exposure, the muscle strips were washed by repeated changes of the Krebs-Ringer buffer at least once every 10 min for 1–2 h before testing Leu-enkephalin. For muscle strips treated with CNA, the concentration of Leu-enkephalin giving 50% inhibition of muscle twitch tension (IC_{50}) was increased from 400 nM (untreated strips) to 2.2 μ M (0.5 nM CNA), 5.5 μ M (3.0 nM CNA), and 173 μ M (25 nM CNA).

CNA is a derivative of the reversible antagonist naltrexone, which, like naloxone, is more potent against enkephalin than against dynorphin_(1–13) in this tissue. For example we found that the apparent K_d of naltrexone is 3 nM against Leu-enkephalin, but 17 nM against dynorphin_(1–13). It was hoped, therefore, that CNA would show a similar differential effect. We exposed the tissue for 20 min to 3 nM CNA then determined the IC_{50} of dynorphin_(1–13), Leu-enkephalin and normorphine. As shown in Fig. 2, CNA treatment alone did not discriminate between the receptors occupied by the three ligands. The IC_{50} values were shifted to about the same degree: 14-, 14- and 12-fold, respectively. As was described previously for Leu-enkephalin, the effect of CNA treatment on dynorphin_(1–13) IC_{50} was not reversed by repeated washing (4.7 nM, 1 h after CNA; 5.3 nM, 5 h after CNA).

To attempt selective protection, we added either the stable enkephalin analogue, DADLE, or dynorphin_(1–13) to the organ bath 1 min before the addition of CNA (3 nM). The lowest concentration of dynorphin_(1–13) able to provide substantial protection of dynorphin_(1–13) potency was 20 nM, and 100 μ M gave complete protection for the 20-min period of exposure to CNA. Comparable concentrations of DADLE were 500 nM and 10 μ M for protection of Leu-enkephalin potency. When muscle strips were incubated with the protective ligand (20 nM dynorphin_(1–13) or 10 μ M DADLE) for 20 min in the absence of CNA and then removed by washing, IC_{50} values were unchanged.

Results of the protection experiments are shown in Fig. 2. When DADLE was present during CNA exposure, the IC_{50} of dynorphin_(1–13) was shifted 17-fold—this was not different from the change in IC_{50} in unprotected muscle strips exposed to CNA alone ($P > 0.05$, Student's *t*-test). However, DADLE significantly reduced the IC_{50} ratio of Leu-enkephalin and of normorphine, from 14 to 3 and 12 to 2, respectively ($P < 0.01$

for both). Thus the protection afforded by DADLE was selective for Leu-enkephalin and normorphine, compared with dynorphin_(1–13). In sharp contrast, when 20 nM dynorphin_(1–13) was used as the protecting ligand, the IC_{50} ratio for dynorphin_(1–13) was reduced significantly, from 14 to 5 ($P < 0.01$), but the IC_{50} ratio for Leu-enkephalin was unaltered. Thus, dynorphin_(1–13) provided selective protection of dynorphin_(1–13) potency.

These results, showing selective protection of Leu-enkephalin and normorphine potency on the one hand, and dynorphin_(1–13) potency on the other, indicate that these ligands act through different binding sites and not through one interconvertible receptor. That DADLE protects both enkephalin and normorphine binding sites is consistent with the hypothesis that these

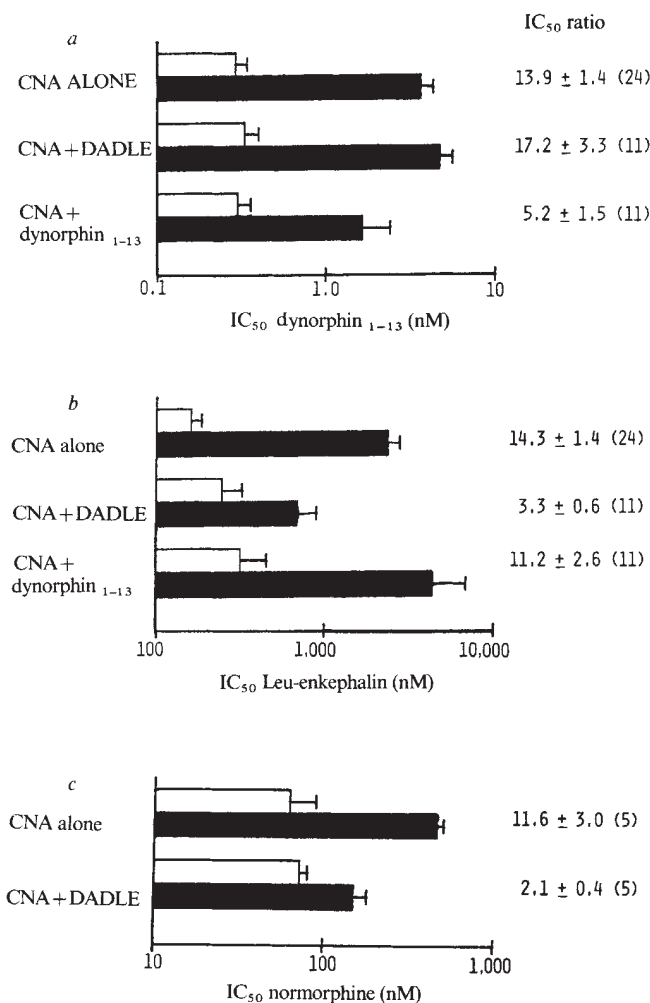


Fig. 2 The effects of CNA treatment on the IC_{50} of a, dynorphin_(1–13), b, Leu-enkephalin, and c, normorphine. For each ligand the IC_{50} was estimated before a 20-min incubation of the muscle strip with 3 nM CNA and again after washing for 1–2 h. Open bars represent mean IC_{50} + s.e.m. before CNA treatment; solid bars represent mean IC_{50} + s.e.m. after treatment of 5–24 muscle strips. Dilutions of dynorphin_(1–13) (Peninsula), Leu-enkephalin (Biosearch), DADLE (Biosearch), and normorphine (Applied Science) were prepared in methanol/0.1 M HCL (1:1, v/v). The IC_{50} was estimated by bracketing the 50% inhibition of the electrically stimulated muscle twitch, and interpolating on a logarithmic concentration-per cent inhibition plot. The protective ligand (either 10 μ M DADLE or 20 nM dynorphin_(1–13)) was added to the bath 1 min before the addition of CNA. Because dynorphin_(1–13) is sensitive to peptidases, after 10 min of the 20-min exposure, the strips were washed once and fresh protective ligand and CNA were added. The IC_{50} ratio is the IC_{50} of the agonist after CNA treatment divided by the IC_{50} before CNA treatment as determined for each muscle strip.

two ligands both act through the μ receptor in the guinea pig myenteric plexus, as suggested by Kosterlitz and Paterson¹⁰.

The paradigm of selective protection has been used before to demonstrate differences in opiate receptor types—Robson and Kosterlitz¹¹ used protection against the alkylating agent phenoxybenzamine and Smith and Simon¹² used protection against *N*-ethylmaleimide to resolve δ and μ binding sites in brain homogenate binding assays. Our study extends that earlier work in two ways. We have used selective protection to distinguish between μ and dynorphin receptors in the guinea pig myenteric plexus and demonstrated that distinction in a functional opiate receptor bioassay system. The presence in guinea pig myenteric plexus of opioid receptors with selective affinity for dynorphin suggests a functional role for this peptide.

We thank Drs P. S. Portoghese and A. E. Takemori for supplying CNA, Drs Brian M. Cox, Francis Leslie, and Carl Romano for helpful discussions, and Pat Lowery for assistance. This work was supported by grants DA-1199 and DA-7063 from the National Institute on Drug Abuse.

Received 27 February; accepted 30 March 1981.

- Goldstein, A., Tachibana, S., Lowney, L. I., Hunkapiller, M. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6666–6670 (1979).
- Chavkin, C. & Goldstein, A. *Soc. Neurosci. Abstr.* **6**, 617 (1980).
- Wuster, M., Schulz, R. & Herz, A. *Neurosci. Lett.* **20**, 79–83 (1980).
- Portoghese, P. S., Larson, D. L., Jiang, J. B., Caruso, T. P. & Takemori, A. E. *J. med. Chem.* **22**, 168–173 (1979).
- Caruso, T. P., Larson, D. L., Portoghese, P. S. & Takemori, A. E. *J. Pharmac. exp. Ther.* **213**, 539–544 (1980).
- Schulz, R. & Goldstein, A. *Br. J. Pharmac.* **48**, 655–666 (1973).
- Goldstein, A., Aronow, L. & Kalman, S. M. *Principles of Drug Action*, 82–104 (Wiley, New York, 1974).
- Rocha e Silva, M. *Eur. J. Pharmac.* **6**, 294–302 (1969).
- Creese, I. & Snyder, S. H. *J. Pharmac. exp. Ther.* **194**, 205–219 (1975).
- Kosterlitz, H. W. & Paterson, S. J. *Proc. R. Soc. B210*, 113–122 (1980).
- Robson, L. E. & Kosterlitz, H. W. *Proc. R. Soc. B205*, 425–432 (1979).
- Smith, J. R. & Simon, E. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 281–284 (1980).

Glutamate-preferring receptors regulate the release of D-[³H]aspartate from rat hippocampal slices

Gethin J. McBean & P. J. Roberts*

Department of Physiology and Pharmacology, School of Biochemical and Physiological Sciences, University of Southampton, Southampton SO9 3TU, UK

At least two of the excitatory afferent pathways in the hippocampal formation, the Schaffer collaterals which extend from the CA3 to CA1 pyramidal cells and the perforant path from the molecular layer of the dentate gyrus, are believed to use glutamate as their transmitter^{1–5}. A third, the commissural pathway, which projects from the CA3/CA4 pyramidal cells to the contralateral CA1 and dentate areas, may be aspartergic⁶. The excitatory transmitter of the remaining afferent pathway, the mossy fibre projection, is as yet unidentified, but there is no evidence that it is either glutamate or aspartate^{6,7}. Despite the obvious importance of excitatory amino acid transmitters in hippocampal function, no evidence of an interaction between these neurotransmitters, such as presynaptic control of release, has been described. Here we report that L-glutamate and several neuroactive analogues inhibit the calcium-dependent, K⁺-evoked release of D-[³H]aspartate from mini-slices of rat hippocampus. This effect was abolished by selective amino acid antagonists and, as it was tetrodotoxin (TTX) insensitive, the effect may be mediated through presynaptic glutamate auto-receptors.

Superfused hippocampal prisms (0.2×0.2 mm), previously loaded with D-[³H]aspartate, attained within 15–20 min a basal release which corresponded to <1% per min of the total accumulated tissue radioactivity (in practice, equivalent to ~3,000 c.p.m. per 10 mg tissue per 5-min perfusion period). The introduction of 50 mM KCl (substituted for sodium) resulted in a >200% increase in the release of the radioactivity. In the presence of a moderately high concentration (100 μ M) of L-glutamate, however, the potassium-evoked release of D-[³H]aspartate was markedly attenuated (Table 1), by ~35%. D-glutamate also inhibited the potassium-stimulated release, as did DL-homocysteate, L-homocysteate and L-cysteate. Note that D-homocysteate, which is a strong excitant of central neurones⁸, was totally ineffective. L-aspartate, in contrast to glutamate, was without effect, as were *N*-methyl-DL-aspartate, ibotenic acid and its derivative, (*RS*)- α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA). Kainic acid did not affect the release of D-[³H]aspartate. None of the agonists tested exerted any inhibitory effect on the spontaneous release of D-[³H]aspartate. This was not unexpected, as for many other systems regulation of transmitter release by presynaptic receptors can only be demonstrated for the calcium-dependent processes of transmitter release, such as occur with depolarization by either potassium or veratridine. For example, the spontaneous release of ³H-GABA (γ -aminobutyric acid) from rat substantia nigra slices was unaffected by agonists, whereas in the presence of 30 mM K⁺, a significant inhibitory effect of GABA was observed⁹.

One possible complicating factor in the study of presynaptic receptors for amino acid transmitters is that of homoeoexchange¹⁰, which would result in an apparent enhancement of spontaneous release by agonists which were substrates for the uptake carrier. Although this phenomenon certainly occurs with synaptosomes, it is probably not evident with slices¹¹ and, certainly in these experiments, enhancement of release was never observed.

Table 1 Effects of excitatory amino acids on the calcium-dependent, K⁺-evoked release of D-[³H] aspartate from hippocampal slices

Agonist	% Inhibition of K ⁺ -stimulated release
L-glutamate (50 μ M)	23 $\pm$ 18
L-glutamate	35 $\pm$ 9*
D-glutamate	32 $\pm$ 19†
L-cysteate	41 $\pm$ 6‡
L-homocysteate	26 $\pm$ 9†
DL-homocysteate	25 $\pm$ 13†
($\pm$) Ibotenate	15 $\pm$ 10
<i>N</i> -methyl-DL-aspartate	—
Kainate	—
AMPA	—
L-aspartate	—

Rat hippocampal tissue prisms (0.2×0.2 mm) were loaded on to gauze rings. One ring was inserted into each of four chambers of a perfusion apparatus²⁰ and a second ring placed above the first. A tightly fitting nylon piston was pressed down on top of the rings. The tissue was superfused at a constant rate of 0.5 ml min⁻¹ with oxygenated Krebs bicarbonate medium (pH 7.4, 37 °C containing 0.5 μ Ci D-[³H]aspartate (specific activity 10 Ci mmol⁻¹) for 15 min to allow high-affinity uptake to occur. Then the slices were rapidly washed and the flow rate increased to 1 ml min⁻¹, and superfusion continued in the absence of label. Released radioactivity was assayed by scintillation counting of 5-min superfusion samples. After achieving a steady basal release rate, 50 mM KCl (substituted for NaCl) was introduced into the perfusion medium for 5 min; this produced a large (213%) calcium-dependent increase in release. L-glutamate (or other analogue) was included with the high-potassium fraction. After stimulation, the medium was exchanged for the low-potassium buffer and a further fraction collected to ensure a return to basal levels. Results are expressed as mean $\pm$ s.e.m. of at least four independent determinations. Significance of difference from control K⁺-evoked release was tested by Student's *t*-test.

**P* < 0.01; †*P* < 0.02; ‡*P* < 0.001.

* To whom correspondence should be addressed.

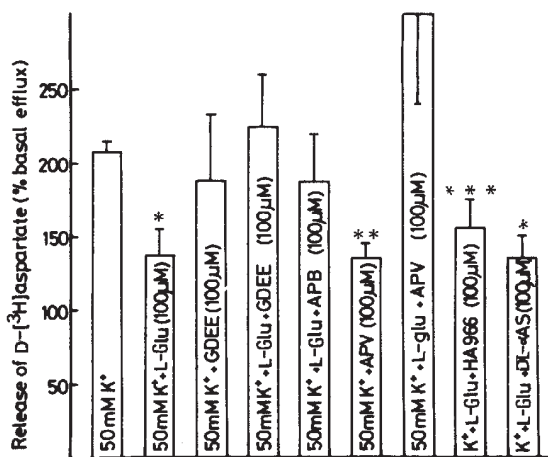


Fig. 1 Effects of amino acid antagonists on the K^+ -stimulated release of D-[3H]aspartate in the presence of 100 μM L-glutamate. Release of D-[3H]aspartate is expressed as a percentage of basal (unstimulated) efflux. Methods are as described in Table 1 legend. Results are mean $\pm$ s.e.m. of at least four determinations. The significance of difference from results obtained with K^+ alone was tested by Student's *t*-test. DL- α AS, DL- α -aminosuberate. * $P < 0.01$; ** $P < 0.001$; *** $P < 0.05$.

The inhibition by 100 μM L-glutamate of the potassium-induced efflux of D-[3H]aspartate could be antagonized by equimolar concentrations of 2-amino-4-phosphonobutyric acid (APB), L-glutamate diethylester (GDEE) and 2-amino-5-phosphonovaleric acid (APV) (Fig. 1). APV alone had a marked glutamate-like action, suggesting that it may act as a partial agonist at these presynaptic receptors. The compounds 1-hydroxy-3-aminopyrrolid-2-one (HA-966) and DL- α -aminosuberate did not affect the inhibition of K^+ -stimulated D-[3H]aspartate release by L-glutamate, and the response was unchanged in the presence of 0.5 μM TTX in the incubation medium.

These results indicate that there are presynaptic receptors in the hippocampus which modulate the potassium-stimulated release of labelled D-aspartate and show a clear preference for compounds known to interact with glutamate receptors. L-aspartate and *N*-methyl aspartate (an aspartate receptor-preferring agonist¹²) were not active. ($\pm$) Ibotenate, which may also interact preferentially with aspartate receptors¹³, was also ineffective in this system. Kainic acid, although it is a potent neuroexcitatory analogue of glutamate, does not seem to interact directly with postsynaptic glutamate receptors¹⁴. Similarly, in this system, there is no evidence for an interaction between kainate and the presynaptic glutamate receptor. With regard to the use of antagonists, it is also significant that HA-966 and DL- α -aminosuberate were inactive as both show marked selectivity for aspartate receptors¹⁵, compared with GDEE, for example, which is strongly glutamate receptor-preferring.

Although D-[3H]aspartate is a valuable tool for studying excitatory amino acid release mechanisms because it is taken up into both glutamatergic and aspartergic terminals as effectively as the natural L-isomers, it is essentially not metabolized by nervous tissue, including the hippocampus¹⁶⁻¹⁸. Thus, although we have identified a presynaptic receptor for glutamate, we cannot say whether this is an autoreceptor or a glutamate receptor which regulates the release of L-aspartate. We have recently begun to investigate the potassium-evoked, Ca^{2+} -dependent release of endogenous L-glutamate from hippocampal slices¹⁹ and have found that transmitter release is again significantly decreased in the presence of glutamate analogues. This implies that, at least in the hippocampus, glutamate regulates its own release through autoreceptors. However, we have been unable to detect a similar system in either striatum or cerebellum.

This work was supported by a SRC grant to P.J.R. We thank Drs J. C. Watkins and P. Krogsgaard-Larsen for gifts of compounds.

Received 21 January; accepted 3 April 1981.

1. Malthe-Sorensen, D., Skrede, K. K. & Fonnum, F. *Neuroscience* **4**, 1255-1263 (1979).
2. Nadler, J. V., White, W. F., Vacca, K., Redburn, D. A. & Cotman, C. W. *J. Neurochem.* **29**, 279-290 (1977).
3. Hamberger, A., Chiang, G., Nylen, E. S., Scheff, S. W. & Cotman, C. W. *Brain Res.* **143**, 549-555 (1978).
4. Hamberger, A., Chiang, G., Nylen, E. S., Scheff, S. W. & Cotman, C. W. *Brain Res.* **148**, 513-530 (1979).
5. Hamberger, A., Chiang, G., Nylen, E. S., Scheff, S. W. & Cotman, C. W. *Brain Res.* **168**, 531-541 (1979).
6. Nadler, J. V., White, W. F., Vacca, K., Perry, B. W. & Cotman, C. W. *J. Neurochem.* **31**, 147-155 (1978).
7. White, W. F., Nadler, J. V. & Cotman, C. W. *Brain Res.* **164**, 177-184 (1979).
8. Balcar, C. J. & Johnston, G. A. R. *J. Neurobiol.* **3**, 295-301 (1972).
9. Kamal, L., Arbilla, S. & Langer, S. Z. in *Presynaptic Receptors* (eds Langer, S. Z., Starke, K. & Dubocovich, M. L.) 193-197 (Pergamon, Oxford, 1979).
10. Levi, M. & Raiteri, M. *Int. Rev. Neurobiol.* **19**, 51-74 (1976).
11. Johnston, G. A. R. in *Glutamate, Transmitter in the Central Nervous System* (eds Roberts, P. J., Storm-Mathisen, J. & Johnston, G. A. R.) 77-87 (Wiley, Chichester, in the press).
12. McCulloch, R. M., Johnston, G. A. R., Game, C. J. A. & Curtis, D. R. *Expl Brain Res.* **21**, 515-518 (1974).
13. McLennan, J. & Lodge, D. *Brain Res.* **169**, 83-90 (1979).
14. Foster, A. C. & Roberts, P. J. *J. Neurochem.* **31**, 1467-1477 (1978).
15. Davies, J. & Watkins, J. C. in *Glutamate as a Neurotransmitter* (eds Di Chiara, G. & Gessa, G. L.) 275-284 (Raven, New York, 1980).
16. Davies, L. P. & Johnston, G. A. R. *J. Neurochem.* **26**, 1007-1010 (1976).
17. Lund Karlson, R. & Fonnum, F. *Brain Res.* **151**, 457-467 (1978).
18. Storm-Mathisen, J. & Woxen Opsahl, M. *Neurosci. Lett.* **9**, 65-70 (1978).
19. McBean, G. J. & Roberts, P. J. (in preparation).
20. Rowlands, G. J. & Roberts, P. J. *Expl Brain Res.* **39**, 239-240 (1980).

Transcription of immunoglobulin heavy-chain sequences from the excluded allele

Phyllis A. Ponte*, Miriam Siekevitz†, Richard C. Schwartz†, Malcolm L. Gefter† & Gail E. Sonenshein*

*Department of Biochemistry, Boston University School of Medicine, Boston, Massachusetts 02118, USA

†Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Production of a functional immunoglobulin heavy-chain gene requires multiple recombination events¹⁻⁵. In germ-line cells, segments of the classical variable portion are encoded in separate genetic elements: variable (V), diversity (D) and joining (J) regions^{3,5,6}. The genes encoding the different clustered constant-region domains (5'- μ , δ , γ^3 , γ^1 , γ^{2b} , γ^{2a} , ϵ , α -3') are widely separated from the variable region⁷⁻¹². During differentiation to immunoglobulin production, one allele containing genes coding for the V, D, J and C segments is rearranged to form a coding sequence for the synthesis of heavy chain. In some cells, the other alleles for heavy-chain segments are rearranged, albeit differently⁸. At the protein level, expression of only one of the alleles is detected (allelic exclusion)^{13,14}. We report here studies on a class of mouse myeloma MOPC 315 variant cell lines which has lost the ability to produce α heavy-chain protein or mRNA. Hybridization analysis indicates that the functional α allele has been lost. However, these variants synthesize unusual RNA species that include α constant-region sequences. We conclude that this transcription is from the 'excluded' allele. Expression of these aberrant RNAs is altered in the presence of a productive gene.

Wild-type mouse myeloma MOPC 315 cells synthesize an immunoglobulin A (IgA, λ^2) protein. We have isolated several variant lines which have lost the ability to produce α heavy-chain protein but have normal levels of λ^2 light chain. To test for the presence of heavy-chain mRNA, cytoplasmic RNA was

translated in a cell-free system and the immunoprecipitated products separated by electrophoresis. RNA from the wild-type cell line (LK) directs the synthesis of both heavy- and light-chain polypeptides (Fig. 1). RNAs from all of the variant lines tested (which includes three additional lines not shown) are unable to direct heavy-chain protein synthesis but normal levels of light chain are produced. Heavy-chain translation at about 1/10th the normal level would have been detected.

To test for the presence of expressed α constant-region sequences, poly(A)-containing cytoplasmic RNA, bound to diazobenzylxymethyl (DBM)-paper¹⁵, was hybridized to ³²P-labelled cloned DNA. With the α probe, RNA from wild-type LK cells shows a strong band of hybridization at ~1,950 bases (Fig. 2a, c). This is the appropriate size for a full-length α heavy-chain mRNA. However, in addition to this mRNA band, there are low levels of smaller molecular weight RNAs of discrete size. To determine whether these RNA species resulted from degradation due to experimental manipulation, the α DNA probe was removed from the DBM filter¹⁵ and the filter rehybridized to λ^2 DNA probe. LK and all the other lines tested

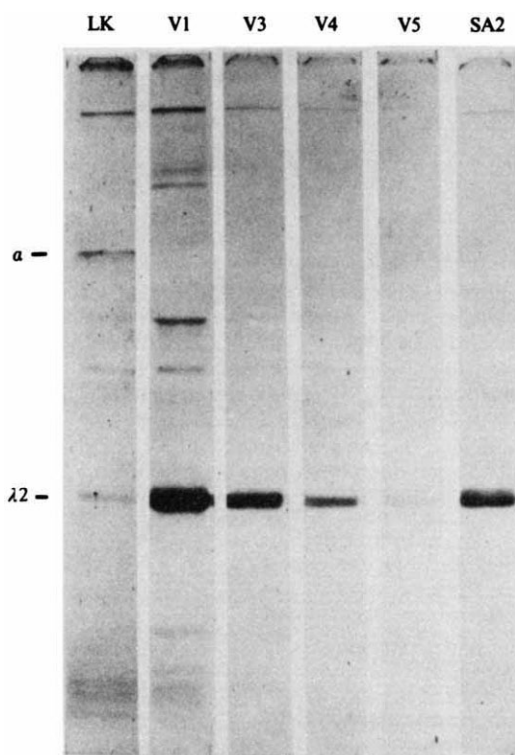


Fig. 1 Cell-free translation analysis of wild-type and variant cell line RNA. MOPC 315 heavy-chain non-producing variant cell lines were isolated from LK (wild type) cells which had been treated with the following mutagens: V1 and V4, ICR 191-OH; V2 and V3, ICR 191-F; V5 was first treated with ethylmethane sulphonate (EMS), and then with ICR 191-F. After mutagenesis, cells were cloned in agar and variant lines isolated using specific antisera to MOPC 315 IgA protein in an overlay assay²⁵. Radioimmunoassays, performed as described elsewhere¹⁶, failed to detect any full-length α heavy-chain protein intracellularly. The assay can detect 10^{-4} times the amount of heavy chain present in steady-state wild-type cells. Cell line SA2 was previously isolated by the same overlay procedure as a spontaneously arising variant of a hybrid cell line derived from fusion of MOPC 315 with MPC 11. The hybrid line (SOO-1) produces all four parent immunoglobulin chains, α , λ^2 , γ^{2b} and κ , whereas SA2 has lost α -chain production but continues to produce the other chains¹⁶. Cytoplasmic RNA was extracted and translated in a rabbit reticulocyte cell free system^{26,27}. The products were precipitated with rabbit anti-315 protein followed by goat anti-rabbit immunoglobulin antisera²⁶. Samples were prepared and electrophoresed on 5–15% SDS-polyacrylamide gels²⁶, which were treated for fluorography²⁸.

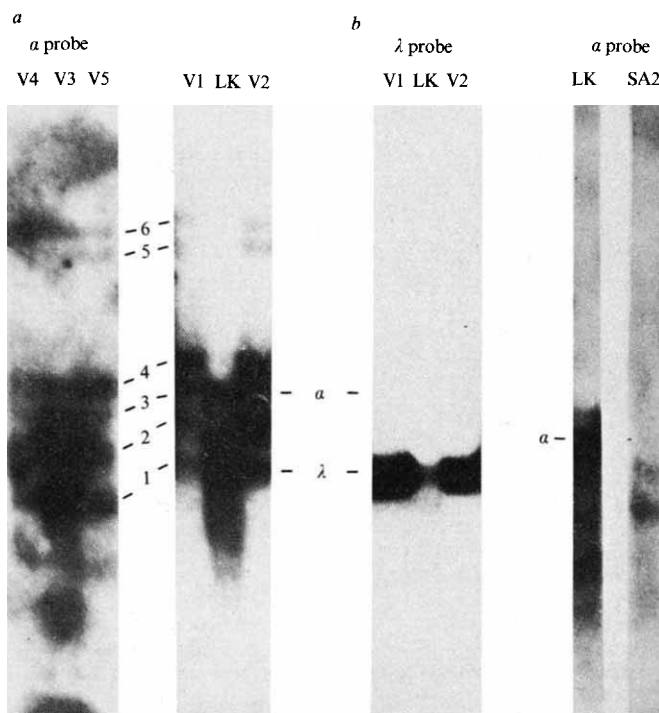


Fig. 2 Comparison of RNA species from wild-type and variant cells. Poly(A)⁺-selected cytoplasmic RNA was denatured with glyoxal²⁹ and electrophoresed on 1.1% agarose gels. The tracks containing LK (wild type) RNA had 2 μ g of RNA, and all those containing variant RNA had 4 μ g. RNA transfer to DBM-paper and hybridization procedures were as described elsewhere¹⁵. The ³²P-labelled probes used were prepared by nick-translation³⁰ to specific activities of $1-2 \times 10^8$ c.p.m. per μ g and $\sim 5 \times 10^6$ c.p.m. were hybridized per track. Position of wild-type full-length heavy- and light-chain mRNAs are indicated as α - and λ -. Visual markers used were rRNA and its precursors prepared from myeloma nuclei. The cloned α heavy-chain DNA used these experiments is 650 base pairs long and corresponds to the 3' end of the heavy-chain mRNA as far as the middle of the constant region. The 870-base pair λ^2 insert is complementary to 3'-end sequences and extends well into the variable region of the light chain. Plasmid pBR322 clones containing inserts of cDNA to MOPC 315 heavy- and light chain mRNA were constructed as described previously²⁰. Cloned DNAs were characterized by hybridization to specific mRNA ('positive selection')³¹ and by hybridization arrest of translation (HART) analysis³².

show a single strong hybridization band at 1,200 bases, the size of λ^2 mRNA (Fig. 2b). Surprisingly, when RNA extracted from five of the independently isolated heavy-chain non-producing variant lines is hybridized to α DNA probe, a complex series of discrete bands is seen (Fig. 2a), none of which seem to co-migrate with α heavy-chain mRNA. The lower bands (1 and 2) apparently correspond to species seen in wild-type cells. Four additional RNA species are also detected; one of these (3) has only a slightly greater mobility than full-length RNA. The other three bands are over full-length size; one is only slightly larger (4) and two are high molecular weight species (5 and 6 are 5,700 and 6,900 bases, respectively). As no nuclear precursor species are detected on hybridization with λ^2 probe (Fig. 2b), it is unlikely that the α RNAs have leaked into the cytoplasm during isolation.

To determine whether the production of aberrant RNA is reflected in the heavy-chain gene organization, we compared the restriction patterns for α -chain coding sequences in wild-type and variant lines. Restriction of wild-type MOPC 315 DNA with *Bam*HI produces two restriction fragments of 12.5 and 9.3 kilobases, containing α constant-region sequences (Fig. 3a); the germ-line restriction pattern gave a single band at 16.5 kilobases⁸. DNA from the heavy-chain non-producing mutant cells

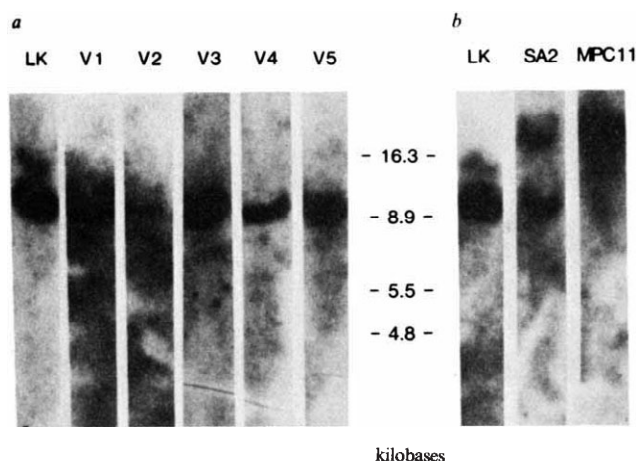


Fig. 3 Hybridization of *Bam*HI-digested DNA to α constant-region sequences. DNA was purified from cells and restricted to completion with *Bam*HI endonuclease. After electrophoresis on a 0.9% agarose gel, the DNA (10 μ g per slot) was transferred to nitrocellulose filters (Millipore) according to Southern³³. Hybridization mixtures contained 5×10^5 c.p.m. per slot of a 32 P-labelled α constant-region probe (2×10^8 c.p.m. per μ g); hybridization was as described elsewhere³⁴. Washes were at 68 °C with $3 \times$ SSC, 0.2% SDS, $10 \times$ Denhardtts for 1 h, $2 \times$ SSC, 0.2% SDS for 1 h followed by 3×1 h with $1 \times$ SSC, 0.2% SDS. ($1 \times$ SSC is 0.15 M NaCl, 0.015 M Na-citrate; $1 \times$ Denhardtts is 0.02% each of Ficoll 400, polyvinyl pyrrolidone and bovine albumin.) Fragment sizes were determined by comparison with λ phage DNA restriction fragments.

have lost the 12.5-kilobase restriction fragment (Fig. 3a). This result is consistent with the 12.5-kilobase band containing the gene coding for the production of heavy-chain mRNA and the 9.3-kilobase band containing the 'allelically excluded' gene. Hybridization with J and switch-region probes has confirmed that the 12.5-kilobase band is the functional allele³⁵. Therefore the aberrant α RNA species in the variants are likely to be due to transcription of the rearranged, but allelically excluded, heavy-chain gene.

Cory and Adams⁸ have shown that the DNA around the α constant region has rearranged in some myelomas producing heavy chains which map upstream to the α constant region, for example, γ^{2b} and γ^1 . Is such a rearrangement of an α heavy chain gene sufficient to promote transcription? MPC 11 (IgG^{2b}), in which one α constant-region allele is rearranged from the germ-line configuration⁸, was tested for the presence of α -specific RNA sequences. No detectable levels of material hybridizing to α constant-region DNA were found in cytoplasmic poly(A)⁺ RNA preparations from this line (data not shown), nor could we detect any transcription of α constant-region sequences in MOPC 21 (IgG1), whose α constant-region genes both remain in the germ-line configuration⁸.

To determine whether the presence of a functional heavy-chain gene other than that for α heavy chain would influence the expression of the excluded allele, RNA from a hybrid line was examined. SA2 is a variant line, isolated from an MPC 11 $\times$ MOPC 315 hybrid cell¹⁶, which has lost the ability to produce α heavy chain (Fig. 1). It synthesizes γ^{2b} heavy chain and a κ light chain from the MPC11 parent and a λ^2 light chain from the MOPC 315 parent.

Cytoplasmic RNA from SA2 contains low levels of the low-molecular weight α constant-region species (Fig. 2c). Interestingly, the RNA species that are larger than full-length message (species 4, 5 and 6) are not detected. Southern hybridization analysis reveals a band at ~16.5 kilobases derived from the MPC 11 parent⁸ and a band at 9.3 kilobases from the MOPC 315 parent (Fig. 3b). The 12.5-kilobase productive α allele is absent. Thus transcription of the 9.3-kilobase allele occurs in the SA2 line. The presence of an active γ^{2b} gene does not 'turn off'

this activity, although the levels of the aberrant RNAs detected are reduced.

Transcription of light-chain gene sequences to give species other than the productive message has recently been described in myeloma cells¹⁷⁻¹⁹. MOPC 315 cells produce a fragment of λ^1 mRNA, as well as full-length λ^2 mRNA^{20,21}. Our results indicate that wild-type MOPC 315 cells synthesize, in addition to a full-length heavy-chain mRNA, lower molecular weight RNA species which contain α constant-region sequences. Similar RNAs are found in variant lines that are unable to produce full-length mRNA due to the absence of the productive allele. Curiously, RNA species of a higher molecular weight than full-length heavy-chain mRNA are also detected in mutant lines. We propose that these species represent transcription from the allelically excluded gene but the nature of these transcripts and their function in the cell is unknown.

Note that the species of greater than full-length message size are only found in cells which produce no heavy-chain protein. They are not detected in the wild-type MOPC 315 or in two other mutant cell lines with altered heavy-chain production (one of which overproduces heavy-chain protein and the other an aberrant heavy chain (manuscript in preparation)). They are also not detected in the hybrid line, SA2, which produces a functional γ^{2b} chain. Although it is not understood how they arise, it may be that the steady-state levels of these aberrant RNAs are regulated by the presence of a functional heavy-chain gene or gene product. This regulation could occur at the level of transcription or RNA splicing. Experiments are being done to determine if these RNA species contain the J region or unspliced α -chain intervening regions²²⁻²⁴. Finally, note that in this case of heavy-chain RNA, and in all the reports of aberrant light-chain transcription¹⁷⁻²¹, allelic exclusion at the level of productive chain synthesis is always maintained.

We thank Dr S. Cory for sharing her results before publication and James McCormack for technical assistance. G.E.S. is an established investigator of the American Heart Association. P.A.P., R.C.S. and M. S. were supported by NIH training grants T32HL07052 and GM 07287. This research was supported by NIH grant AI16051 and American Cancer Society grant IM-214. All experiments involving recombinant DNA were conducted in accordance with NIH guidelines.

Received 9 December 1980; accepted 25 March 1981.

- Dreyer, W. J. & Bennett, J. C. *Proc. natn. Acad. Sci. U.S.A.* **54**, 864-869 (1965).
- Sakano, H. *et al. Nature* **277**, 627-633 (1979).
- Early, P. W., Davis, M. M., Kaback, D. B., Davidson, N. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 857-861 (1979).
- Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. *Cell* **19**, 981-992 (1980).
- Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676-683 (1980).
- Schilling, J., Clevinger, B., Davie, J. M. & Hood, L. *Nature* **283**, 35-40 (1980).
- Honjo, T. & Kataoka, T. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2140-2144 (1978).
- Cory, S. & Adams, J. M. *Cell* **19**, 37-51 (1980).
- Cory, S., Jackson, J. & Adams, J. M. *Nature* **285**, 450-456 (1980).
- Shimizu, A. *et al. Nature* **289**, 149-153 (1981).
- Rabbitts, T. H., Forster, A., Dunnick, W. & Bentley, D. L. *Nature* **283**, 351-356 (1980).
- Yaoita, Y. & Honjo, T. *Nature* **286**, 850-853 (1980).
- Pernis, B., Chiappino, G., Kelus, A. S. & Gell, P. J. *exp. Med.* **122**, 853-875 (1965).
- Cebra, J. S., Colberg, J. E. & Dray, S. J. *exp. Med.* **123**, 547-558 (1966).
- Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350-5354 (1977).
- Siebert, G. R., Harris, J. F. & Gefter, M. L. *J. Immun.* **121**, 1808-1818 (1978).
- Seidman, J. G. & Leder, P. *Nature* **276**, 790-795 (1978).
- Perry, R. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1937-1941 (1980).
- Alt, F. W., Enea, V., Bothwell, A. L. M. & Baltimore, D. *Cell* **21**, 1-12 (1980).
- Bothwell, A. *et al. Nature* **290**, 65-67 (1981).
- Schwartz, R. C., Sonenshein, G. E., Bothwell, A. & Gefter, M. L. *J. Immun.* (in the press).
- Davis, M. M. *et al. Nature* **283**, 733-739 (1980).
- Gilmore-Hebert, M. & Wall, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 342-345 (1978).
- Schibler, U., Marcu, K. B. & Perry, R. P. *Cell* **15**, 1495-1509 (1978).
- Coffino, P. & Scharff, M. *Proc. natn. Acad. Sci. U.S.A.* **68**, 219-223 (1971).
- Sonenshein, G. E., Siekevitz, M., Siebert, G. R. & Gefter, M. J. *exp. Med.* **148**, 301-312 (1978).
- Pelham, H. R. & Jackson, R. J. *Eur. J. Biochem.* **67**, 247-256 (1976).
- Bonner, W. M. & Laskey, K. A. *Eur. J. Biochem.* **46**, 83-88 (1974).
- McMaster, G. K. & Carmichael, G. C. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4835-4838 (1977).
- Rigby, P. W. J., Diekmann, M., Rhodes, C. & Berg, P. J. *molec. Biol.* **113**, 237-251 (1977).
- Ricciardi, R., Miller, J. & Roberts, B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4927-4931 (1979).
- Paterson, B. M., Roberts, B. E. & Kuff, E. L. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4370-4374 (1977).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Jeffreys, A. J. & Flavell, R. A. *Cell* **12**, 429-439 (1977).
- Cory, S., Webb, J. & Adams, J. *Biochemistry* (in the press).

BOOK REVIEWS

Why do as you would be done by?

Vernon Reynolds

MERE mention of the word "ethics" tends to make people's eyes go glassy. Yet, well trodden as it is, the field of ethical philosophy is wide open to new ideas. Indeed, Singer believes we are still only at the "beginning of a long march", only now obtaining the "knowledge [that] may clear up the confusion that surrounds ethics" (p.168).

The new knowledge that Singer injects into ethical philosophy in this book is the theory and related facts of biological altruism. Without defining ethics as such, he treats it for the most part (except in a section towards the end) as a quality of personal action, that quality which is concerned not so much with what we do, or why we do it, as with our evaluation of it, whether it is right or wrong, good or bad. As he shows, with great perseverance and lucidity, this evaluation is independent of the actions themselves, and therefore no amount of objective demonstration of the causes of our acting in particular ways can explain ethics.

But Singer is aware that this tends to leave ethics "hanging in the air" as a rootless *a priori*, and this, he feels, is a weakness. In what, if anything, might ethics be rooted? In religion? In social rules? In rational thought? In biology? All have a share, but the deepest roots are in biology. The theory of kin and reciprocal altruism, while wholly unable to give form or content to ethical evaluations, can, he believes, help us to understand the human being as an ethical animal, by showing that many species have, for genetic reasons, evolved patterns of behaviour which foster the survival and reproductive success of others in their own kin group or social group. To this extent "ethics is part of the natural human condition" (p.23).

Singer denies that there is any general difference between human beings and animals. We are all part of a common evolutionary process. We all share "common elements" (p.29) and these have a biological basis. What these elements are is not specified, but genes are frequently referred to, and Singer is clear that mankind cannot escape from the process of natural selection. He is equally clear, though, that in the case of such a complex information-processing and often creatively thinking organism as *Homo sapiens*, the appeal to genes as such for the explanation of right-wrong evaluations is falsely based. What then is the relation between our ethics and biological altruism?

First we should look at Singer's biology. He outlines his theory of altruism in Chapter 1. It is gene-based, and adds up

The Expanding Circle: Ethics and Sociobiology. By Peter Singer. Pp.190. ISBN 0-374-15112-1. (Farrar, Straus & Giroux: 1981.) \$10.95. To be published in Europe and the Commonwealth on July 30th by Oxford University Press, £6.95.

the combined positive effects of an organism's behaviour on the survival-to-reproductive-age chances of its kin and neighbours to establish the probability of transmission. Here he goes, rightly, beyond the Lorenzian concern with motivation-structure (aggression, fear, sex) to find the roots of altruism not in the outcome of motivated decisions about behaviour but in the inevitable results of the genetic process, the blind "logic of living systems" as François Jacob calls it. Wilson, Dawkins, Trivers, Hamilton, Smith, Prince, Parker, all have their share in the ideas expressed.

Of these, Wilson gets the lion's share of attention. He comes under fire from Singer in his claim that a biology of ethics will do away with the need for an independent ethics (and ethical philosophers), being able to settle matters of ethics by reference to predictable biological outcomes. This, for Singer, overstates the place of biology in relation to ethics. Values come first, and then the predicted outcomes of actions and all other relevant knowledge can be assessed, and plans made and acted on. Singer does not accept the use of scientific theories or data as a basis for an ethics of the here-and-now; this for him is a case of the naturalistic fallacy, moving from "is" to "ought", a logically indefensible step first shown to be such by Hume in 1739. Whereas science can provide explanations and facts, ethics consists of directives. Facts have no power to direct; they are neutral; ethics has power to direct but cannot provide facts or explanations. We need ethics and facts for rational planning at all levels — personal, group, race or species.

The matter of levels clearly concerns Singer, perhaps because of the emphasis in evolutionary biology on the individual or the gene as the normal unit of selection. In striving for a unified theory that will link here-and-now ethics to its biological base, there is not just the problem of priorities discussed above, but the problem of how in a process based on differentials between individuals there can arise an ethics transcending individual interests and reaching to those of the group, the species and even — remember that Singer was author of *Animal Liberation* (Cape, 1976) — of other species. This is the problem of

the expanding circle: how, he wants to know, has ethics come, in many cultures and at various times, to transcend itself, from personal to tribal considerations, from tribal to racial, from racial to mankind, from mankind to all sentient beings? Are all such extensions outgrowths of the sentiments evolved as a result of kin and reciprocal altruism? And if so, what is the mechanism of the process?

Singer argues that the mechanism is the application of rational thought and action, which can transcend the situations provided by the kin-group and co-residential group. The process is the growth of universal values out of small-group values. The pay-offs, though not specified by Singer, are presumably pay-offs at the whole-group, national or even wider levels, in terms of political stability and economic security. The rational element he stresses is the exhortation "do as you would be done by". On this basis, people act not only for themselves, or their kin. They also act for their groups, their nationals everywhere, the members of other races, all members of the species and even members of other species involved in the common life-process on our planet (other planets aren't considered!). All can be brought under a set of ethical principles; indeed, he sees this as "progressive" (p.117).

But because of the expansion of ethics in this way we should not suppose that one or two clear-cut directives will suffice; that would be abandoning rational ethics for an ethics based on intuition or ancestral custom or some other principle. Singer is not in favour of such a rigid ethics; he wants a flexible, situation-specific ethics that can take account of the complexity of the world. Just as Mary Midgley, searching for the common good, concluded that *bonum est multiplex*, Singer concludes that the rules of ethics are at all times "subject to critical scrutiny" (p.164).

The book is a short one. Because of the felicitous style, it is an easy read. But it is packed with ideas needing further exploration.

In particular, the ethics of "do as you would be done by" has to be squared with the existence of groups who would be done by in different ways. One is reminded, at the individual level, of the dilemma faced by the young queer from Khartoum, who took a lesbian up to his room. . . . On the wider level, do we do unto Nazis, knowing their ideological bent, as we would be done by? Singer, to be sure, wants flexibility. But how much, when and where? Has he really validated his title? Are there not

times when the circle must be contracted, to nation, family or self? I think Singer will readily agree that there are, but he has not pursued this aspect as fully as it deserves; one could start, perhaps, from the work of Coser on the functions of social conflict, or Evans Pritchard's analysis of segmentary social systems.

Second, anthropologists must surely find it hard to accept the approach to "rationality" adopted by Singer. He speaks of rationality as if it were a unitary phenomenon, leading properly, given a particular set of considerations, to certain conclusions. Not only is this all too rarely the case within cultures sharing one set of concepts — one language etc. — but between cultures the idea of what it is to

think or behave rationally varies in many dimensions. It is doubtful indeed whether all cultures even have a notion of rationality. Who knows but it may be advantageous, in order to overcome the "cognitive dissonance" problem he refers to (p.143), to dispense with it completely?

Criticisms aside, however, this book is undoubtedly an excellent exposition of a most complex problem. Peter Singer has without question clarified the many subtleties and singularities involved in the relation between ethics and socio-biology. □

Vernon Reynolds is University Lecturer in Physical Anthropology at the University of Oxford, and author of The Biology of Human Action (W.H. Freeman, 2nd Edn, 1980).

Thus armed, Richardson reviews the experimental literature on imagery and memory. Chapters 4 and 5 are on imagery and immediate memory and on pictorial imagery, Chapters 6–8 deal primarily with imagery and long-term memory, and the last chapter concerns individual differences. He argues that whereas the effects of imagery on long-term memory for single words and sentences are best explained by a propositional theory, the use of imagery in immediate memory (such as rotating or inspecting a figure mentally) and in remembering pictures is best explained in terms of a dual-coding theory. This suggests that mental images may have different functional properties in different instances and, consequently, the evidence from these instances should be weighed separately in deciding theoretical issues.

I regret that he chose not to discuss experiments in the phenomenology of imagery (how big is an image, how wide an angle does it subtend, how long does it take to conjure it up?), since it would have fitted nicely into the chapter on imagery and working memory. I also wish that he had dealt more fully with neuropsychological approaches to problems in mental imagery. Recent investigations by Bisiach and Luzzatti (*Cortex* 14, 129–133; 1978), who found that individuals with left-sided neglect also neglect the left side of their mental image, suggest that such investigations have a useful contribution to make.

However, the book is an excellent and very readable source of information on research in mental imagery. Richardson has acted as referee and thereby clarified and made more interesting an area that was badly in need of organization. It is left to another book to offer new directions to research and theory □

Morris Moscovitch is an Associate Professor in the Psychology Department and Centre for Research in Human Development at Erindale College, University of Toronto.

Imagination more firmly controlled

Morris Moscovitch

Mental Imagery and Human Memory. By J.T.E. Richardson. Pp.178. ISBN 0-333-25370-1/0-312-52975-9. (Macmillan, London/St Martin's Press: 1980.) £16, \$22.50.

AFTER being neglected for about a half-century by experimental psychologists in the English-speaking world, mental imagery once again become a worthy subject of inquiry in the middle 1960s. The fact that such private, mental representations could be studied objectively signalled the end of the old stimulus-response behaviourism and augured the birth of a new cognitive psychology that could explore freely a variety of phenomena that are part of mental life.

By the mid-1970s, however, the new cognitive psychology had become the established doctrine and research on mental imagery had grown duller while, at the same time, the controversy surrounding it had become more intense. Currently, much research on mental imagery seems to be concerned with establishing the nature of the code in which mental images are represented rather than in exploring further the nature of the phenomenon or the uses to which it can be put. One group believes that mental images are represented in a code that is fundamentally different from that in which verbal information is represented. The former is concrete, analogical and perhaps isomorphic with perception, whereas the latter is abstract and propositional. Against this dual-coding position is the view that all information is represented in terms of a common code involving networks of propositions. The controversy has taken on the trappings of an arms race. No sooner has one side unveiled a powerful empirical or theoretical weapon which threatens the other with ultimate defeat, than an equally destructive finding or argument is developed by the other side. One can only lament this expenditure of

talent, because the differences between the two camps, though large in theory, have become almost indistinguishable in practice. It often takes more than an expert to decide whether an experiment struck a blow for one side or the other.

Such expert refereeing is provided by J. T. E. Richardson in *Mental Imagery and Human Memory*. As the title suggests, he is concerned with mental imagery only insofar as it contributes to human memory. His aims, though limited, are not modest. They are to analyse and comment on conceptual and methodological problems and to integrate and organize the research findings in an area that has grown exponentially over the years and has become quite unwieldy as a result. By and large he has done a good job.

In the early chapters he reviews critically, yet succinctly, the major historical trends in imagery research that led to the current state of affairs. He suggests that some of the conceptual problems facing imagery research arose in part as a result of making imagery acceptable to a behaviouristically orientated psychology. By providing a conceptual analysis of mental imagery as it is used in ordinary language and psychological discourse, Richardson shows up the shortcomings of some of the current approaches and clarifies the theoretical arguments between the single- and dual-coding theorists. Particularly important here is the idea that mental images, because they are representations, have the logical feature of intentionality. That is, the object and functional origin of the image is determined by how it is intended, not by what it resembles. Thus, images arise from "a system of knowledge which is conceptual and propositional" (p.40). Once constructed, however, images may have emergent properties which could not be computed from the original descriptions and which may not be coded in propositional form.

Estuarine mixture

J. D. Burton

Chemistry and Biogeochemistry of Estuaries. Edited by E. Olausson and I. Cato. Pp.452. ISBN 0-471-27679-0. (Wiley-Interscience: 1980.) £28.50, \$78.40.

THE study of estuaries presents a notable challenge to aquatic geochemists because of the range of physicochemical conditions and variations in composition encountered in estuarine waters and sediments, and the variability and complexity of estuarine environments. There have been several reasons for taking up this challenge and for the consequent emergence, during the past decade, of estuarine chemistry as a distinctive field. The complex pathways of

chemical species in estuaries provide important examples of the operation of many geochemical processes. Knowledge of the source and sink functions of estuaries is crucial to an understanding of the chemical mass balances in the shelf seas and oceans. In addition to these fundamental aspects, the extensive — and sometimes conflicting — uses of many estuaries for industrial, commercial and recreational purposes, have led to a need, in terms of resource management, to understand the estuarine behaviour of pollutants.

The need for multidisciplinary awareness in the study of environmental chemical processes is particularly apparent in estuarine chemistry, and the present book appropriately contains four chapters which deal with geomorphology and classification of estuaries, physical processes of water transport and mixing, and sedimentary and biological processes. They serve very well the purpose of providing background, and will also be consulted in their own right. The wisdom of treating the chemistry in as many as nine further chapters may be questioned. Despite the degree of subdivision — which means that the emerging coherence of the subject tends to be obscured — there remain significant omissions and imbalances which are aggravated by the problem, common to many edited works, of uneven treatment. Thus, although there are two contributions dealing with specific groups of organic pollutants (halogenated hydrocarbons and petroleum hydrocarbons), the more general organic chemistry of estuaries, including the role of humic materials, is largely overlooked. The use of radionuclides in the natural decay series to study estuarine processes is scarcely discussed, certainly not commensurately with the importance of the insights which have been obtained in this way. The carbon dioxide system, a key topic, is accorded a separate chapter but nevertheless receives remarkably cursory treatment.

The strengths lie mainly in the greater than usual account taken of the extensive European literature on the subject, and in the high quality of some individual chapters. Among these, there is a lucid treatment of equilibrium modelling of chemical speciation by Dyrssen and Wedborg, a well integrated and well documented discussion by Duinker of suspended particles and their interactions, and a welcome account by Presley and Trefry of interstitial waters of sediments. An unusual and interesting topic is the recovery and decontamination of polluted estuaries, discussed by Cato, Olsson and Rosenberg, with particular emphasis on benthic communities. The strong points are sufficient to make the book a useful addition to the literature of estuarine science and one which should interest a wider audience than just the specialists in estuarine chemistry.

J.D. Burton is Reader in Chemical Oceanography at the University of Southampton.

Hominid heroes through the kaleidoscope

Andrew Hill

Missing Links: The Hunt for Earliest Man. By J. Reader. Pp.272. ISBN 0-00-216091-9/0-361-73590-6. (Collins/Little, Brown: 1981.) £9.95, \$19.95.
Lucy: The Beginnings of Humankind. By D.C. Johanson and M.A. Edey. Pp.409. ISBN 0-671-25036-1/0-246-11362-6. (Simon and Schuster/Granada: 1981.) \$16.95, £9.95.

PICTURE yourself in a camp by a river, in the remote Afar desert of northern Ethiopia. One of the oldest and most spectacular of fossil hominids has just been found; Beatles' music plays loudly over the expedition tape recorder; the fossil — parts of a female skeleton — becomes affectionately known as "Lucy" after *Lucy in the Sky with Diamonds*. The book of the same name, co-authored by Don Johanson, one of the expedition's leaders, is an extravagant romp across the slippery surface of palaeoanthropology. It tells the story of the International Afar Research Expedition and its work, with a bit of palaeoanthropology's historical background, a bit of Johanson's historical background, a bit of gossip. The story is an exciting one, for the expedition has collected large numbers of hominid specimens, including Lucy, which are over three million years old, and probably the most important recently found. Along with similar fossils from Tanzania they are the oldest known hominids from anywhere in the world, and the collection represents a considerable achievement. With a little help from his friends Johanson interprets it in a way he describes as controversial, by suggesting quite plausibly that it all comes from one species, which he has named *Australopithecus afarensis*. This species was dentally and cranially primitive, but was fully bipedal, and he suggests it gave rise not only to the rest of the plio-pleistocene australopithecines, but is also the ancestor of that other prominent hominid lineage — *Homo*, leading to ourselves.

With this material it would be difficult to write a totally uninteresting book, but nevertheless I was disappointed, although initially well disposed to it. Most recent popular books on the subject have come from one factory, and it might have been rewarding to sample another brand. Also the promise of a subjective account was attractive; one that treated the people behind the palaeontology, their personalities, motives and rivalries. After all palaeoanthropology has always been one of those activities where not uncommonly participants gently stab each other in the front, and some consideration of this might have revealed useful information about the dynamics of the science as well as successfully appealing to the general reader. Unfortunately most of the gossip

seems gratuitous, or at least simply contrived to reinforce Johanson's ego and status in the discipline, which gradually emerges as one of his chief concerns. Whilst initially amusing, it eventually becomes a little tedious. Similarly the style, with its lapses into chatty dialogue, does not work for me, and at times has a patronizing aspect. Others, however, may find the book a good read, and as it has been well promoted it may do good by interesting the public in a fascinating field. A personalized, subjective account depends largely for its effect on how interesting the author is. Despite its title the real hero of the tale is not Lucy, nor even palaeoanthropology, but Don Johanson himself. If you are interested in Don Johanson, this is the book for you.

If you are interested in palaeoanthropology the book for you is *Missing Links* by John Reader. Unlike Johanson, Reader has no pretensions to being one of the world's leading palaeoanthropologists. He is not a professional scientist, but a very professional photojournalist with a varied career in Africa and elsewhere. His previous book concerned African wildlife; now we have Reader's digest of fossil man.

But this is a serious book and deserves serious treatment. It is the best popular account of palaeoanthropology I have ever read, and simultaneously makes an original and stimulating contribution to the science. Perhaps its most striking feature is Reader's outstanding photographs of fossils, discoverers and memorabilia, but his text matches these. It provides a palaeo-anthropological history in terms of its major discoverers and discoveries from 1857 to 1980. This information, which alone is intrinsically interesting, is framed by a theme now common enough in the exegesis of scientific method but which is seldom publicized. Reader indicates that hypotheses are prior to data, and that they are heavily conditioned by social factors, preconceived notions and prejudice. He suggests the acceptability of ideas is closely connected with "... the personality and persuasive ability of their proposer. . .", that theories "... reveal as much of the

Journals supplement

On October 1 *Nature* will publish a review supplement devoted to new journals. The supplement will cover science journals which have as their initial publication date the period between January 1978 and May 1980, which appear at least three times a year and which use English as the primary language. Publishers and societies are invited to submit four sample issues, including the first and most recent, of journals which fit the above criteria for consideration for review. Journals should be sent to Reviews Editor, *Nature*, 4 Little Essex St, London WC2R 3LF, England.

scientist as of the science ...". Palaeoanthropology is an eminently suitable field in which to pursue this idea, a field guaranteed to give Feyerabend a good time. Many of its foremost practitioners include people of strong personality and persuasive ability, often well equipped with preconceived notions. Also, like a number of sciences with an historical element, the constraints upon conjecture are relatively slight, for the material evidence for any particular theory is rare, and even more rarely unequivocal. "Controversial" simply means that other people do not believe what you say, and this condition is not new in palaeoanthropology. Each scientist's beliefs can often be justified scientifically, depending for example upon his opinion about time, extent of variation, sexual dimorphism, or the weight given to various anatomical characters. Reader traces these kinds of influences over the past hundred years or so with considerable

insight, consideration and knowledge. He does not ridicule people or their beliefs and attitudes, merely suggests that this is the way the science has advanced; and advanced it has despite these strong personal elements in its methodology.

It would be a great pity if *Missing Links* fell into the shadow of other books by perhaps better known or better publicized authors, for it deserves a wide readership. Not only is it an entertaining and informative book for the general public, but it should not be ignored as an anthropology textbook, providing a thoroughly documented history from a viewpoint that more palaeoanthropologists would do well to share. They might remember the girl with kaleidoscope eyes — just shake the evidence and it falls into another pattern, often equally attractive as the one before.

Andrew Hill is a Research Fellow at the National Museums of Kenya and at Yale University.

Understanding the facts of IR spectroscopy

Richard C. Lord

Advances in Infrared Group Frequencies. The Infrared Spectra of Complex Molecules, Vol.2, 2nd Edn. By L.J. Bellamy. Pp.299. ISBN 0-412-221350-3. (Chapman & Hall: 1981.) £15, \$35.

In the past ten years or so there has been a relative decline in the use of infrared spectroscopy for the solution of problems in structural organic chemistry, primarily because of the widening scope of nuclear magnetic resonance and other techniques. For lack of practitioners among recent graduates, skilful interpretation of the infrared spectrum of a complex molecule is becoming something of a lost art. In the preface to this book, the author puts the reason for this in a nutshell: "One cannot interpret an infrared spectrum without a good knowledge of the experimental facts, but too rigid an interpretation without any understanding of how these facts originate can lead to gross errors". Happily, he has provided us with an admirable summary of the experimental facts in successive editions of *The Infrared Spectra of Complex Molecules*. Now, with the second edition of Vol.2 of this work, he gives an up-to-date and comprehensive basis for an "understanding of how these facts originate".

The book is quite compact. The eight chapters cover less than 300 pages and treat successively the alkanes, carbon-carbon double bonds, triple bonds and cumulative double bonds, X-H stretching frequencies, carbonyl frequencies, other double-bond vibrations, the stretching frequencies of XO₂ groups and the effects of hydrogen bonding on the infrared

spectrum. Each chapter provides a thorough discussion of mechanical, geometrical and electronic factors that affect the group frequencies under consideration. There are up-to-date references (to 1979) at the end of each chapter, totalling about 1,500 in all.

To me, the chapters on olefinic double-bond frequencies, carbonyl vibrations and hydrogen bonding seem particularly useful. That on carbonyl frequencies is rightly the most extensive; a band due to a carbonyl group is often the most intense in the spectrum of an organic molecule and its precise frequency is a measure of the net effect of a multitude of factors. These include the masses of adjacent atoms, their geometry, interaction with neighbouring vibrations or with overtones (Fermi resonance), conjugation with double bonds and other electronic effects, and intermolecular forces. After detailed discussion of these factors, they are summarized by a table of some 500 carbonyl frequencies in suitably illustrative organic molecules.

This volume gives a thorough account of the various origins of group frequencies in organic molecules. It is a valuable work of reference for those who wish to improve their talents in the interpretation of the infrared spectra of such molecules. Taken with the companion Vol.1, it gives a comprehensive survey of the state of our understanding of the infrared spectra of complex molecules. □

Richard C. Lord is the Director Emeritus of the MIT Spectroscopy Laboratory, where he founded in 1950 the first postgraduate training course in the technique and applications of infrared spectroscopy.

Plastid development

F. R. Whatley

Chloroplasts. Results and Problems in Cell Differentiation, 10. Edited by J. Reinert. Pp.240. ISBN 3-540-10082-2. (Springer-Verlag: 1980.) DM 78, \$46.10.

THIS book can be read in sequence from cover to cover and represents a beautifully balanced treatment of plastid development and differentiation. All the articles are pitched at the same level and it is easy to cross-refer from one to the other.

In the opening contribution, Schnepf lays down the framework for the rest of the book in describing the development and interconversions of plastids; using evidence from electron micrographs, he provides a classification of all known types. Butterfass next discusses evidence for the continuity of plastids, and also the coordination of the plastid population, both in terms of the origin of the organelles by division and the evolution of patterns of plastid maintenance.

At a more biochemical level, R. G. Herrmann and Possingham give an account of plastid DNA, its amount and arrangement in the circular chromosome; Wollgiehn and Parthier put a similar emphasis on biochemistry in their discussion of RNA in plastids and aspects of protein synthesis. The important chlorophyll-protein complexes associated with light harvesting are then treated in some detail by F. H. Herrmann *et al.* in the chapter on thylakoid biosynthesis, which also includes a discussion of the sites of synthesis of particular components under the control of nuclear and plastid DNA.

Fraction I protein (ribulosebiphosphate carboxylase) is discussed by Bottomley, who considers the composition and origin of the component sub-units, and Sundqvist *et al.* rehearse the influence of the main external factors on chloroplast differentiation, especially the effects of light, hormones, minerals and water stress. Finally, the experimental evidence that chloroplasts can be cultured outside the cell is drawn together by Leech, who points out the possibilities of experiments which can be done with surviving plastids.

Regrettably, many of the articles end with a note, added in proof, that the literature for this review was completed in 1976. Since the book was published some four years later the editors must have experienced serious delays in receiving some of the manuscripts. In spite of this, all concerned are to be congratulated on the impressively balanced treatment, which covers structural, physiological and biochemical considerations. As well as being of great help to research workers, advanced undergraduates should also benefit from the book. □

F. R. Whatley is Sherardian Professor of Botany at the University of Oxford.